Oceanographic modeling and data assimilation: Adapting tools to support MRV for CDR

Matthew Long
Climate & Global Dynamics Laboratory
National Center for Atmospheric Research

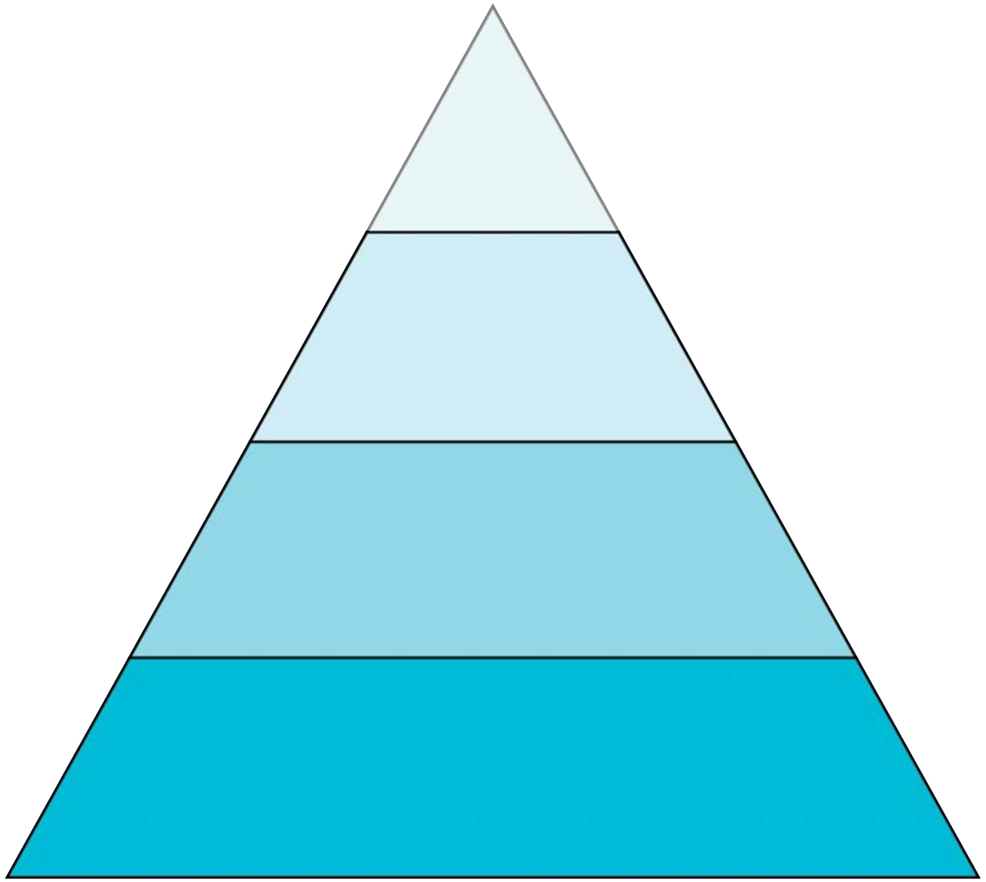
NCAR

16 June 2022

NSF

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

Modeling frameworks to support ocean CDR & MRV



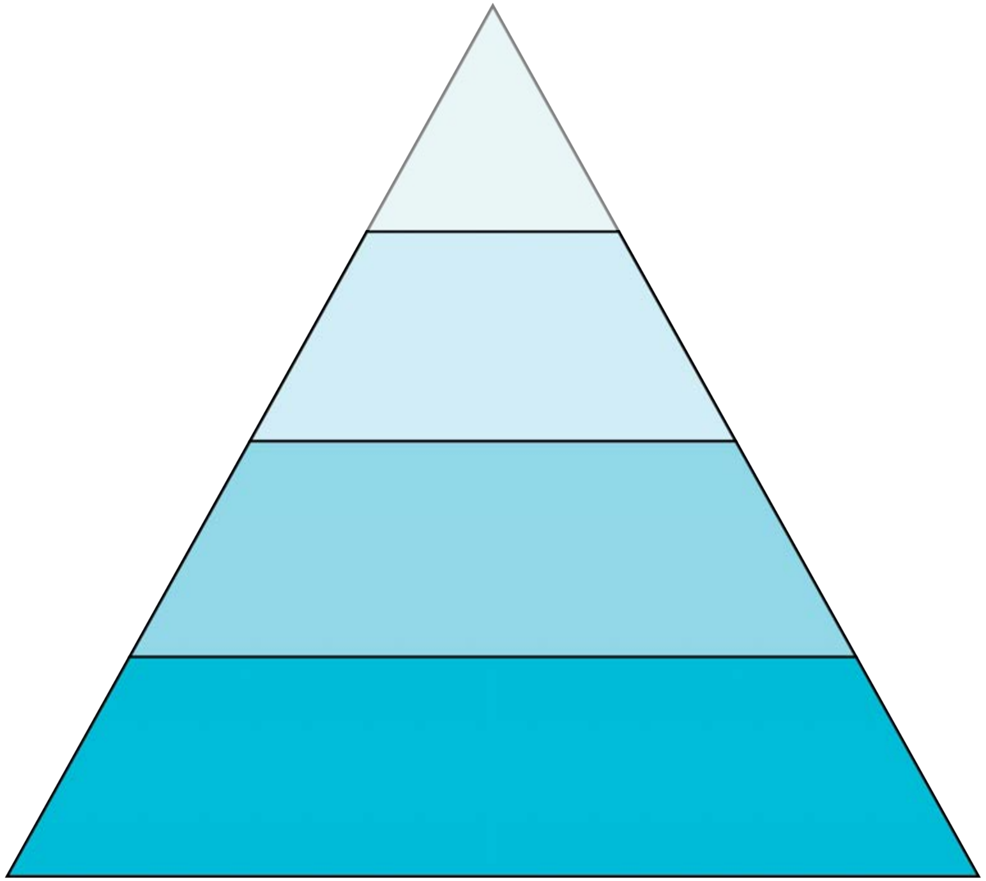
CDR processes & experimental framework

Observing system design
Verification with sequential data assimilation

Data assimilation

Ocean physical and biogeochemical models

Modeling frameworks to support ocean CDR & MRV



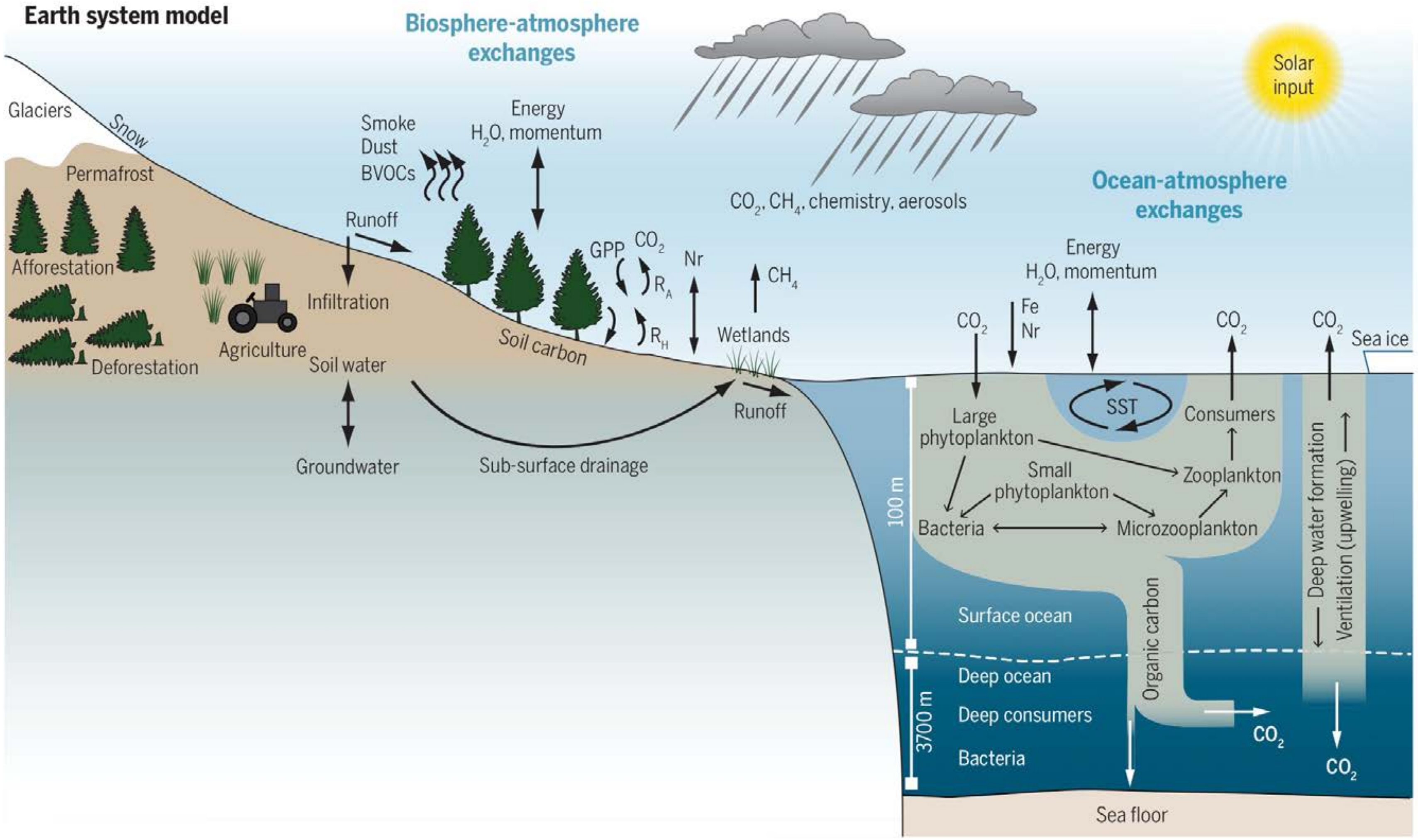
CDR processes & experimental framework

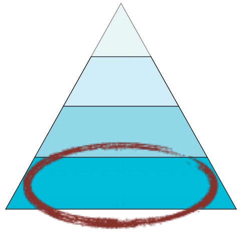
Observing system design
Verification with sequential data assimilation

Data assimilation

Ocean physical and biogeochemical models

Earth system models



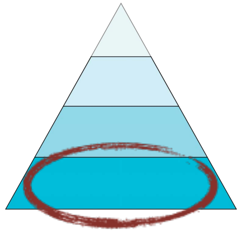


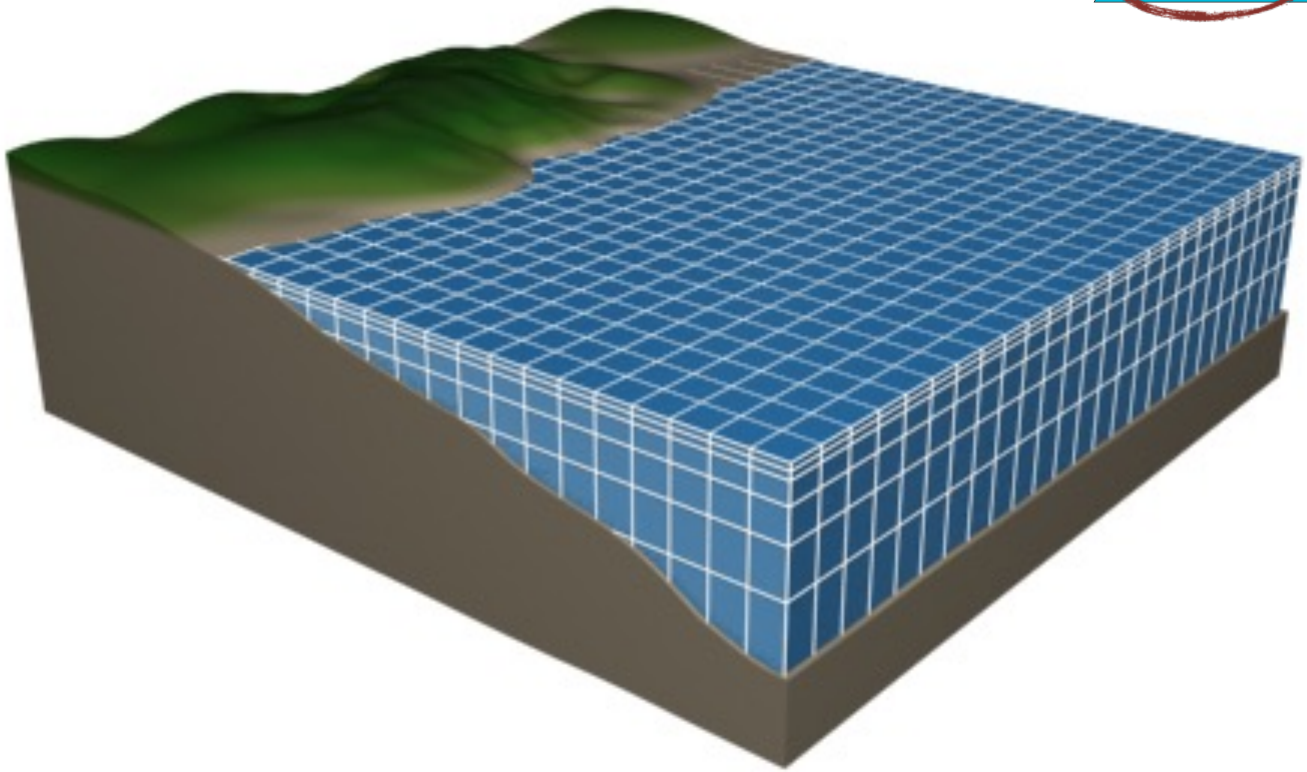
Bonan & Doney, *Science* (2019)

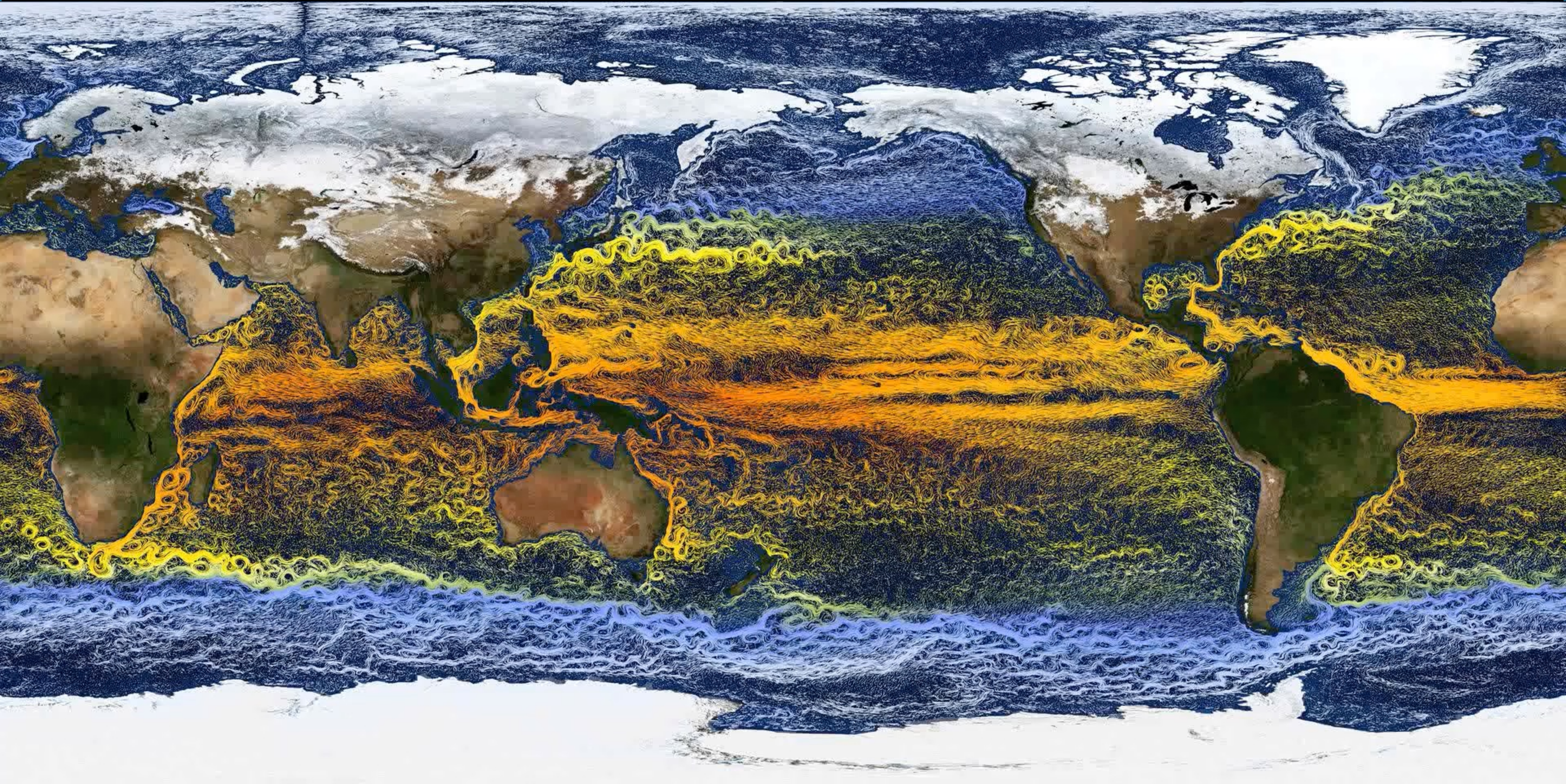
General circulation models

Navier-Stokes: Equations are “exact”, but cannot be solve analytically

General Circulation Models (GCMs) implement numerical solutions on discrete grid

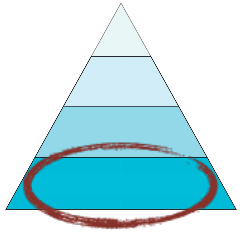






https://ihesp.github.io/archive

General circulation models

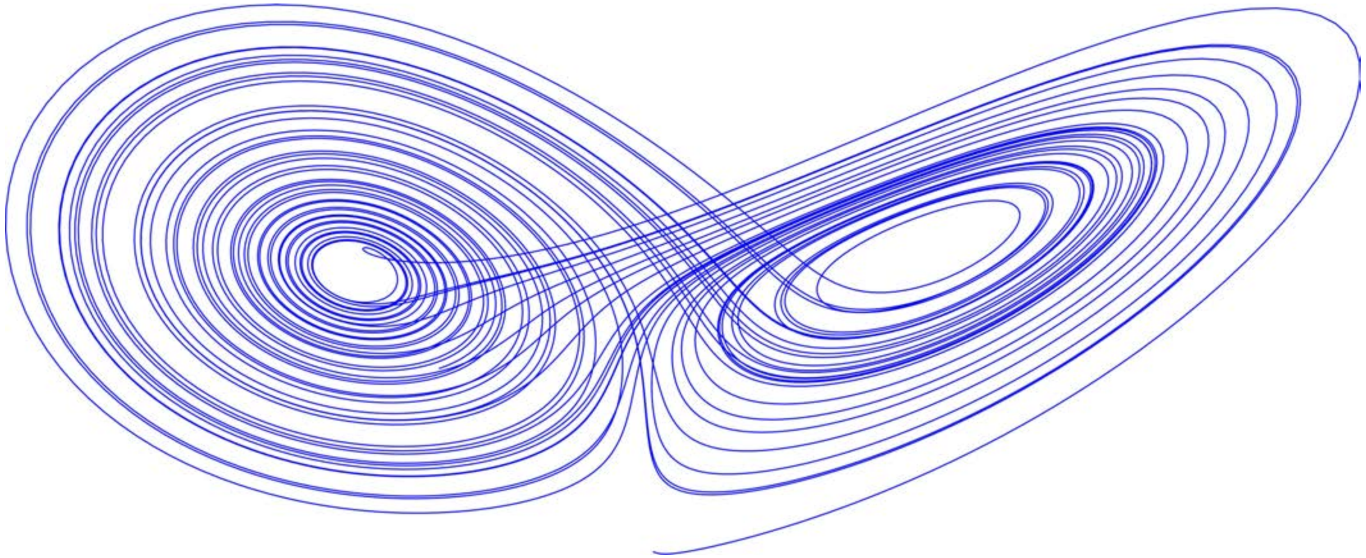


Navier-Stokes: Equations are “exact”,
but cannot be solve analytically

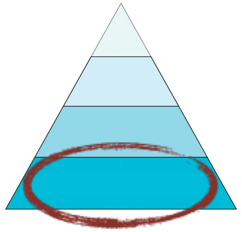
General Circulation Models (GCMs) implement numerical solutions on discrete grid

Deterministic, but chaotic:

- Perpetual novelty
- Infinite sensitivity to initial conditions



General circulation models

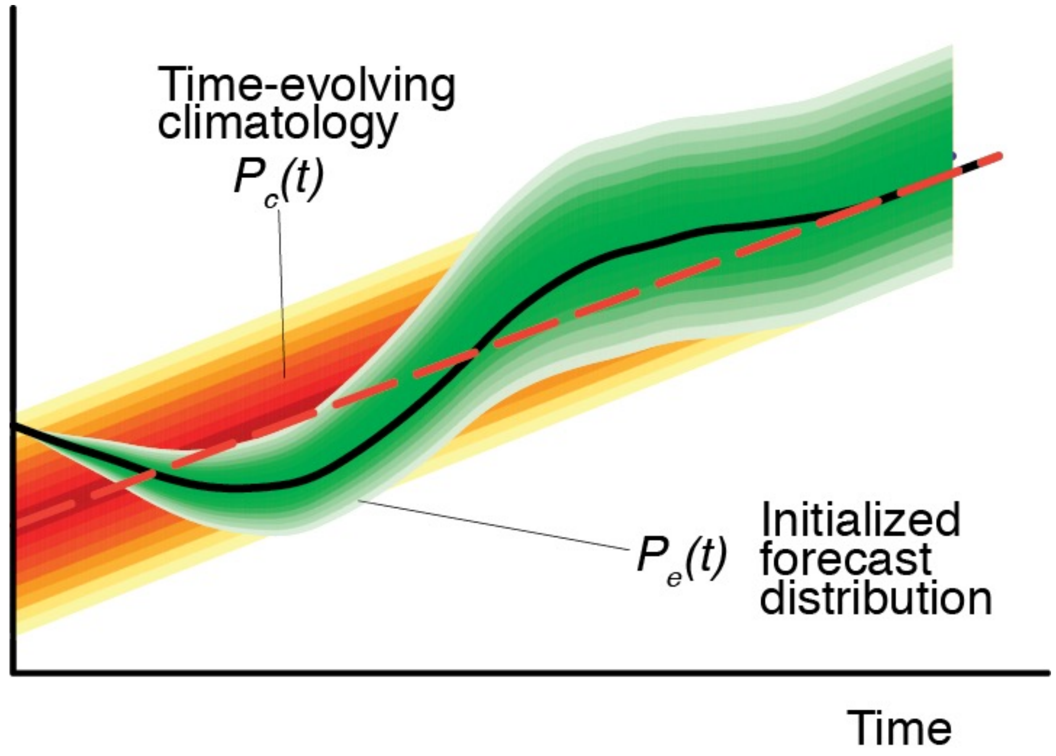


Navier-Stokes: Equations are “exact”, but cannot be solve analytically

General Circulation Models (GCMs) implement numerical solutions on discrete grid

Deterministic, but chaotic:

- Perpetual novelty
- Infinite sensitivity to initial conditions



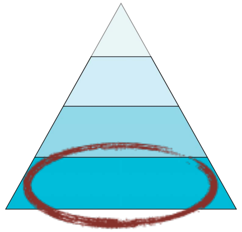
General circulation models

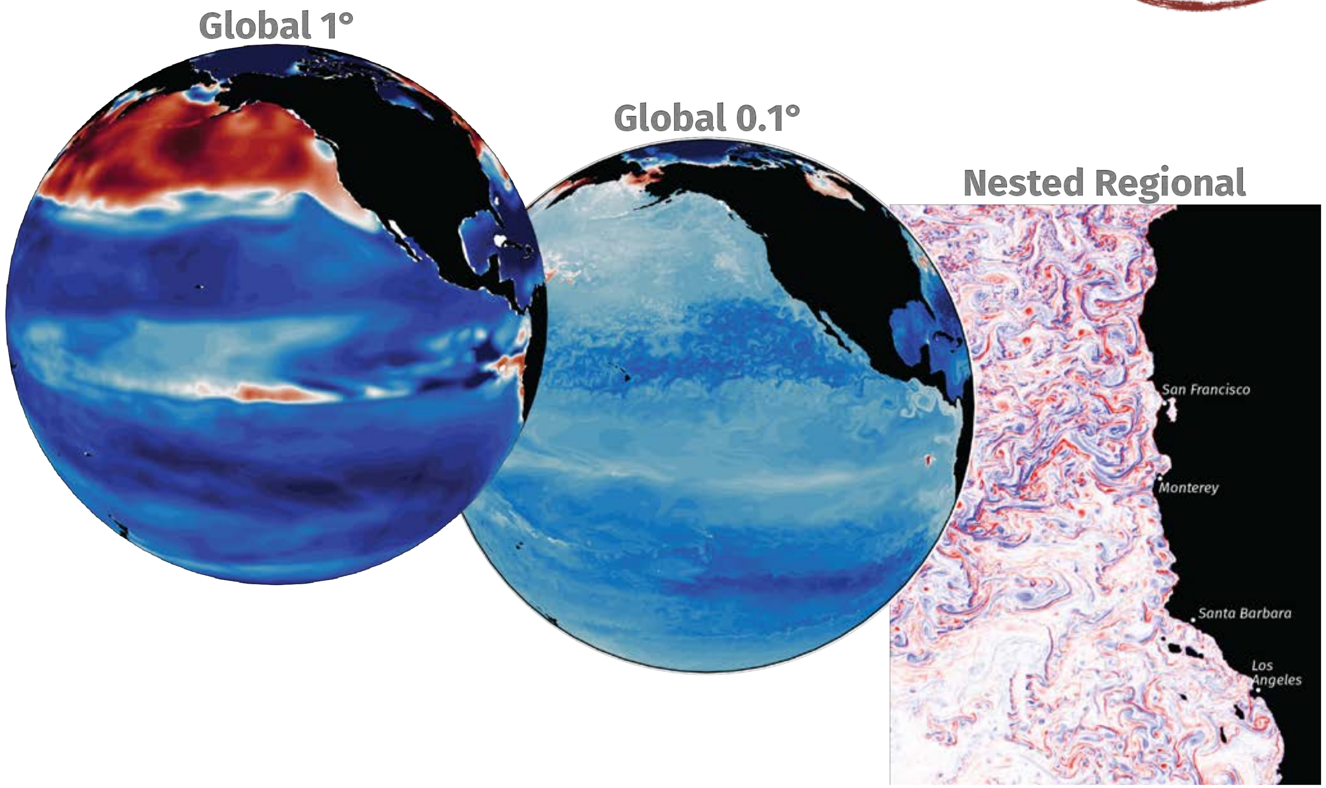
Navier-Stokes: Equations are “exact”, but cannot be solve analytically

Primitive equation models:
- Numerical solutions on discrete grid
- General Circulation Models (GCMs)

Deterministic, but chaotic:
- Perpetual novelty
- Infinite sensitivity to initial conditions

Scale interactions: energy cascade requires parameterization of subgrid-scale effects

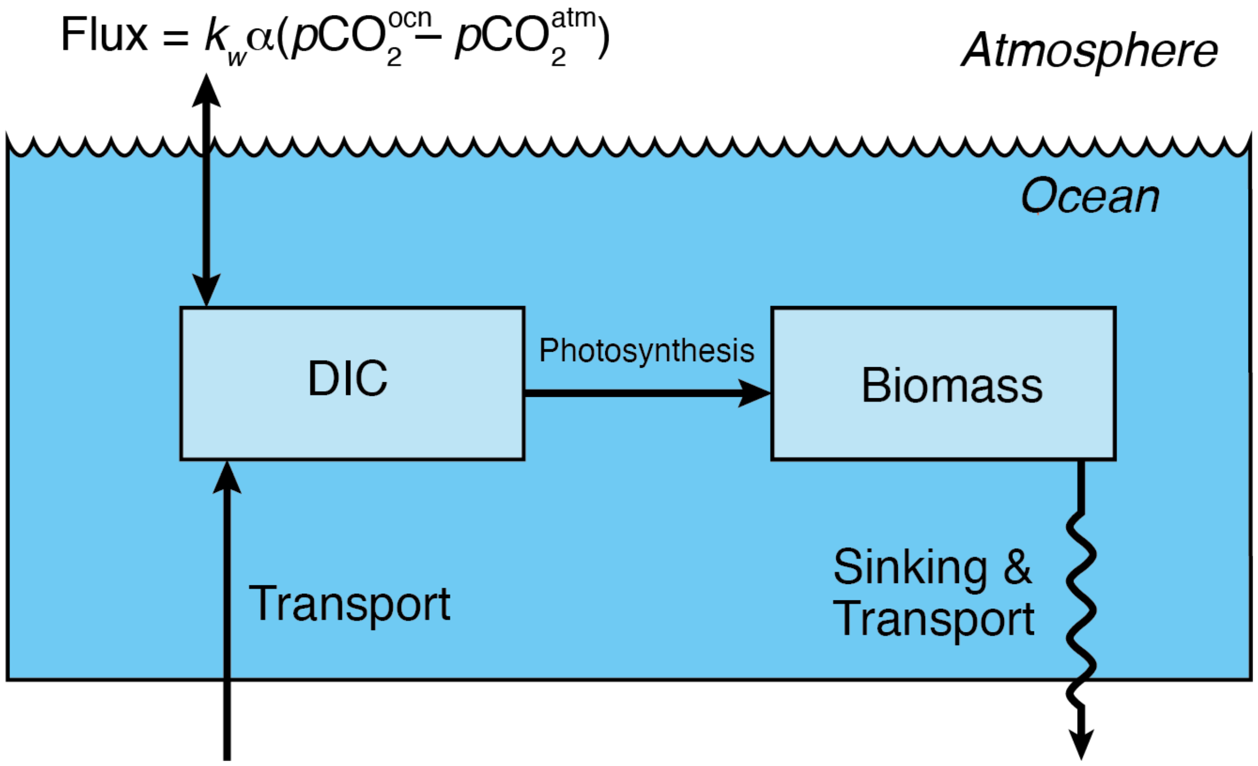




Courtesy of S. Bachman (NCAR)

Representing ocean carbon biogeochemical dynamics

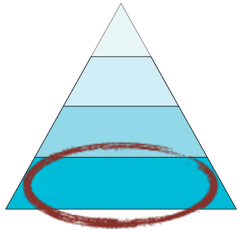
Biology sets upper ocean DIC budget



Dissolved inorganic carbon

$$DIC = \underset{\sim 0.5\%}{[H_2CO_3^*]} + \underset{\sim 88.6\%}{[HCO_3^-]} + \underset{\sim 10.9\%}{[CO_3^{2-}]}$$

Representing ocean carbon biogeochemical dynamics

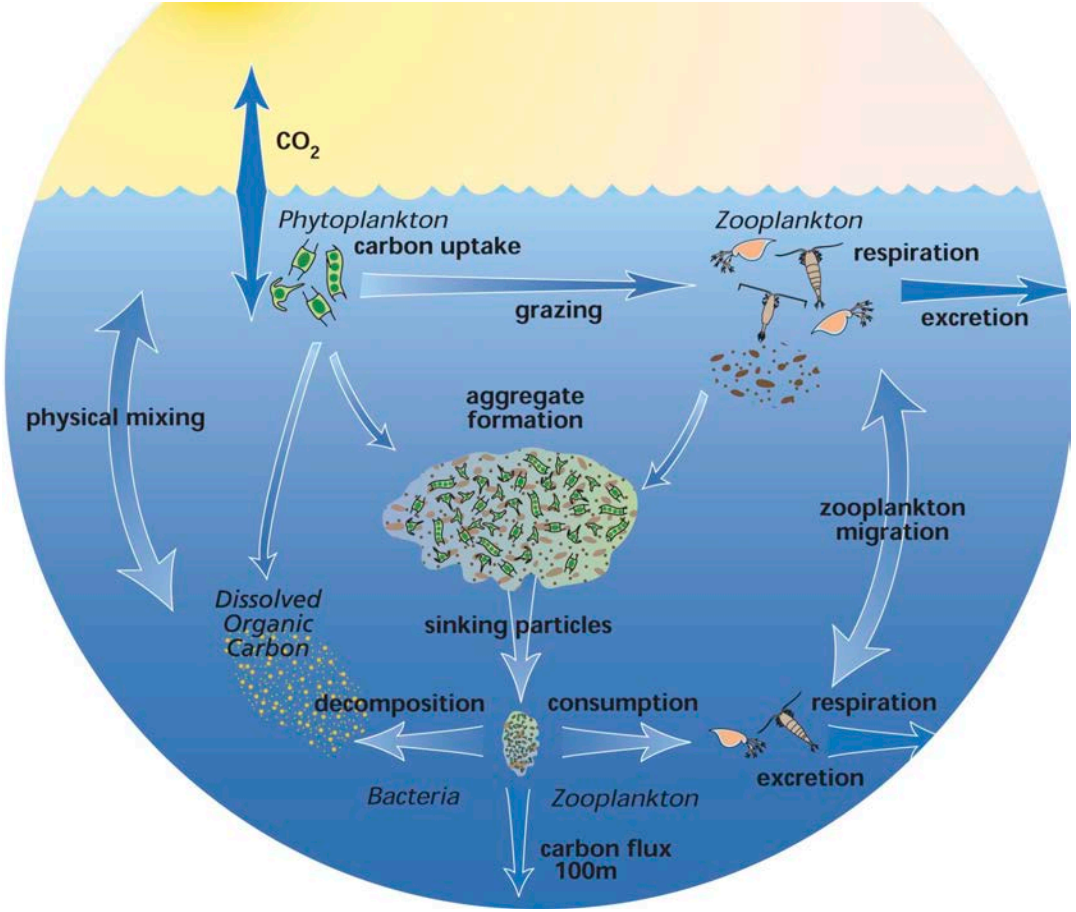


Biology sets upper ocean DIC budget

Ecosystem function controls **C export**

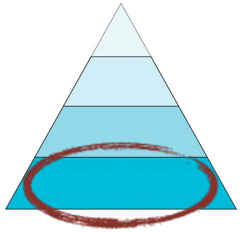
There are no universally-accepted governing equations:

- Plankton function types
- Transfer functions



US-JGOFS

Representing ocean carbon biogeochemical dynamics



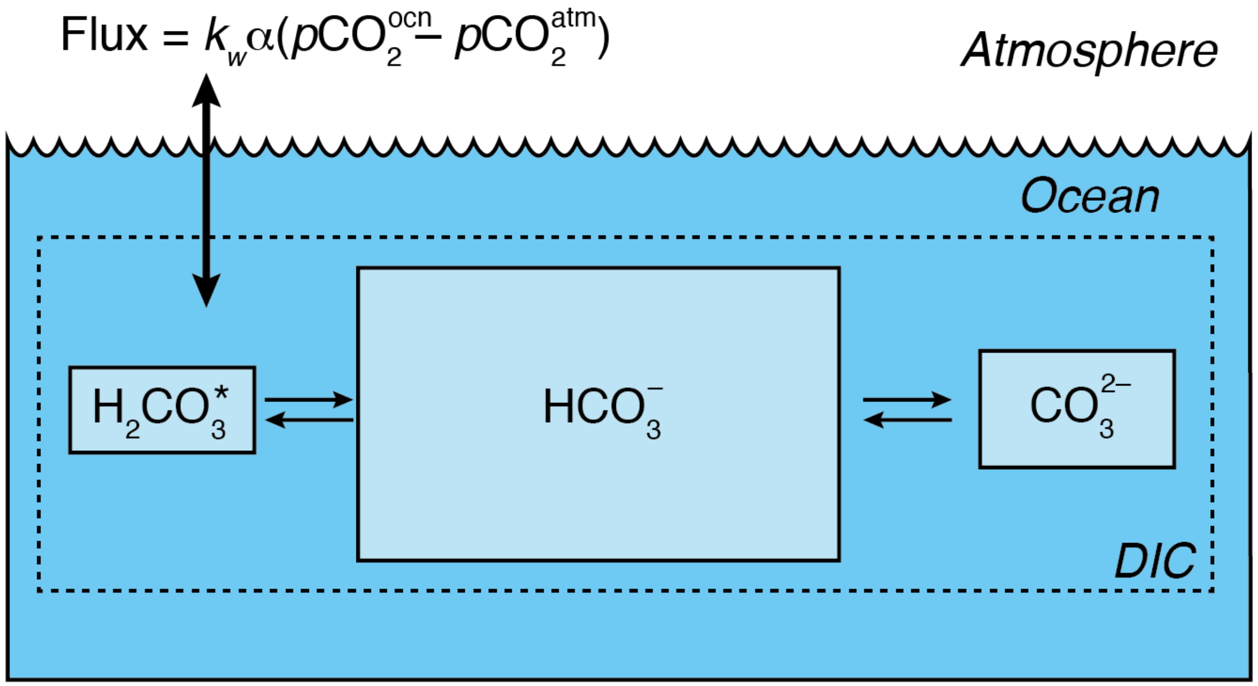
Biology sets upper ocean DIC budget

Ecosystem function controls **C export**

There are no universally-accepted governing equations:

- Plankton function types
- Transfer functions

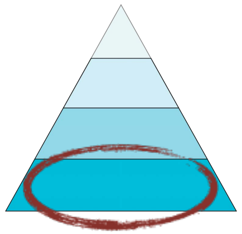
Explicit carbon chemistry



Dissolved inorganic carbon

$$DIC = \underbrace{[H_2CO_3^*]}_{\sim 0.5\%} + \underbrace{[HCO_3^-]}_{\sim 88.6\%} + \underbrace{[CO_3^{2-}]}_{\sim 10.9\%}$$

Representing ocean carbon biogeochemical dynamics



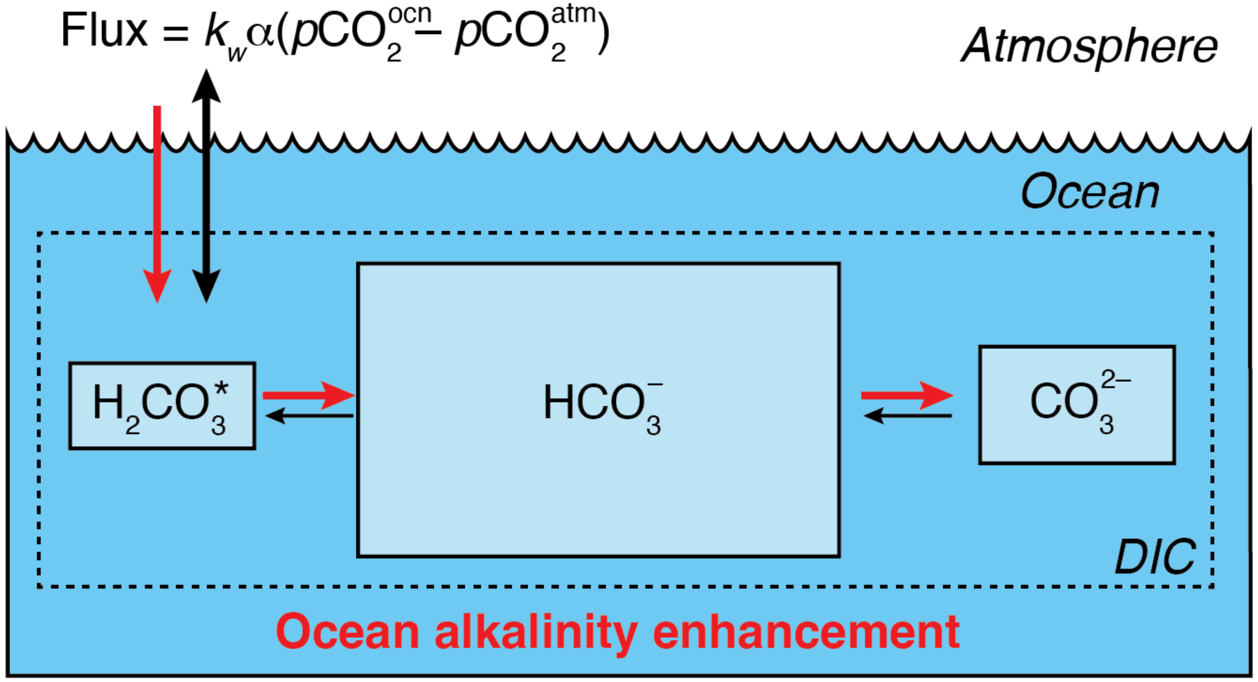
Biology sets upper ocean DIC budget

Ecosystem function controls **C export**

There are no universally-accepted governing equations:

- Plankton function types
- Transfer functions

Explicit carbon chemistry



Dissolved inorganic carbon

$$DIC = \underset{\sim 0.5\%}{[H_2CO_3^*]} + \underset{\sim 88.6\%}{[HCO_3^-]} + \underset{\sim 10.9\%}{[CO_3^{2-}]}$$

Representing ocean carbon biogeochemical dynamics

Biology sets upper ocean DIC budget

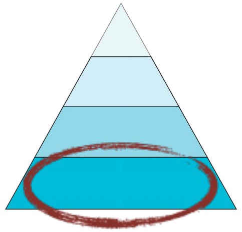
Ecosystem function controls **C export**

There are no universally-accepted governing equations:

- Plankton function types
- Transfer functions

Explicit carbon chemistry

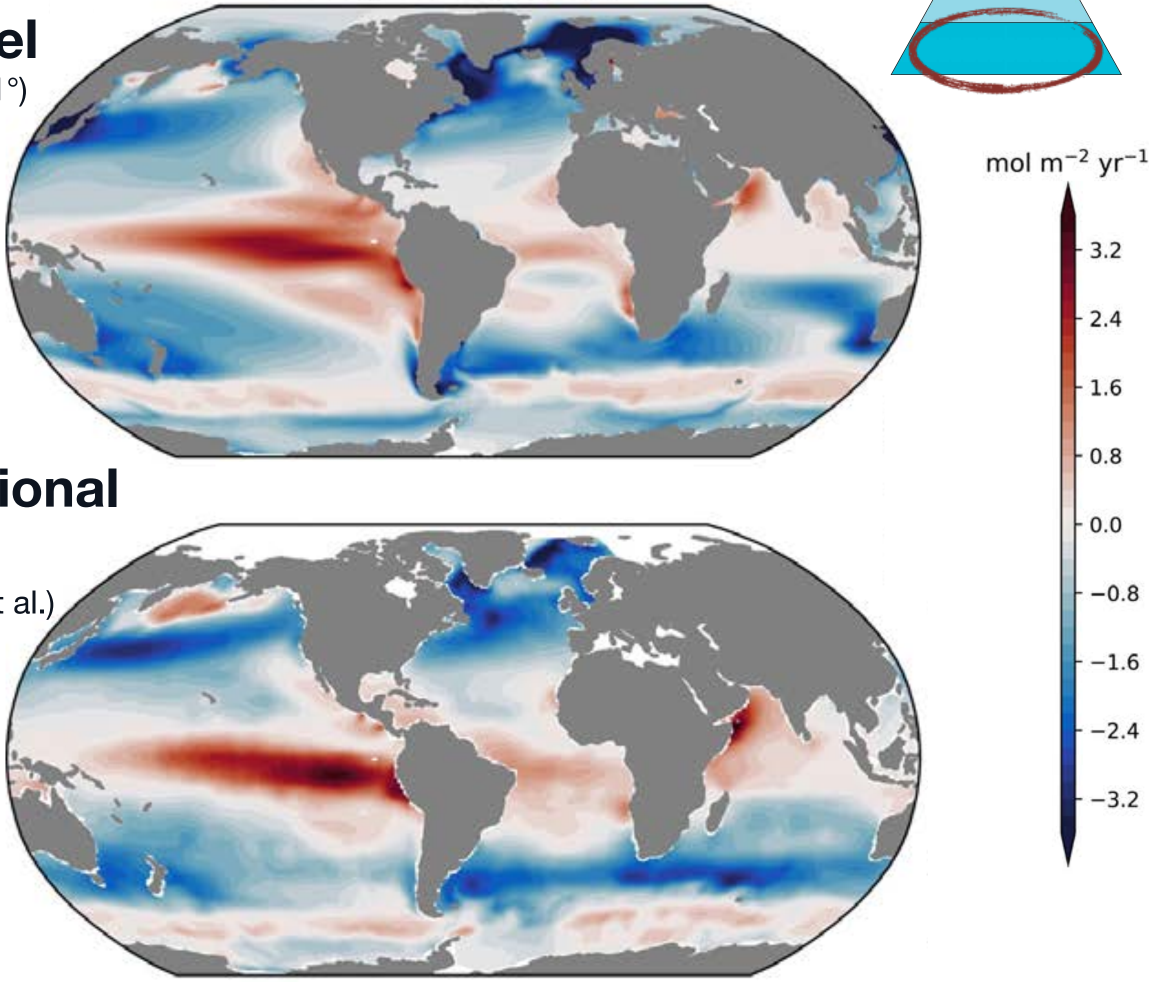
Models are built at the process-level; simulations show emergent features



Air-sea CO_2 flux

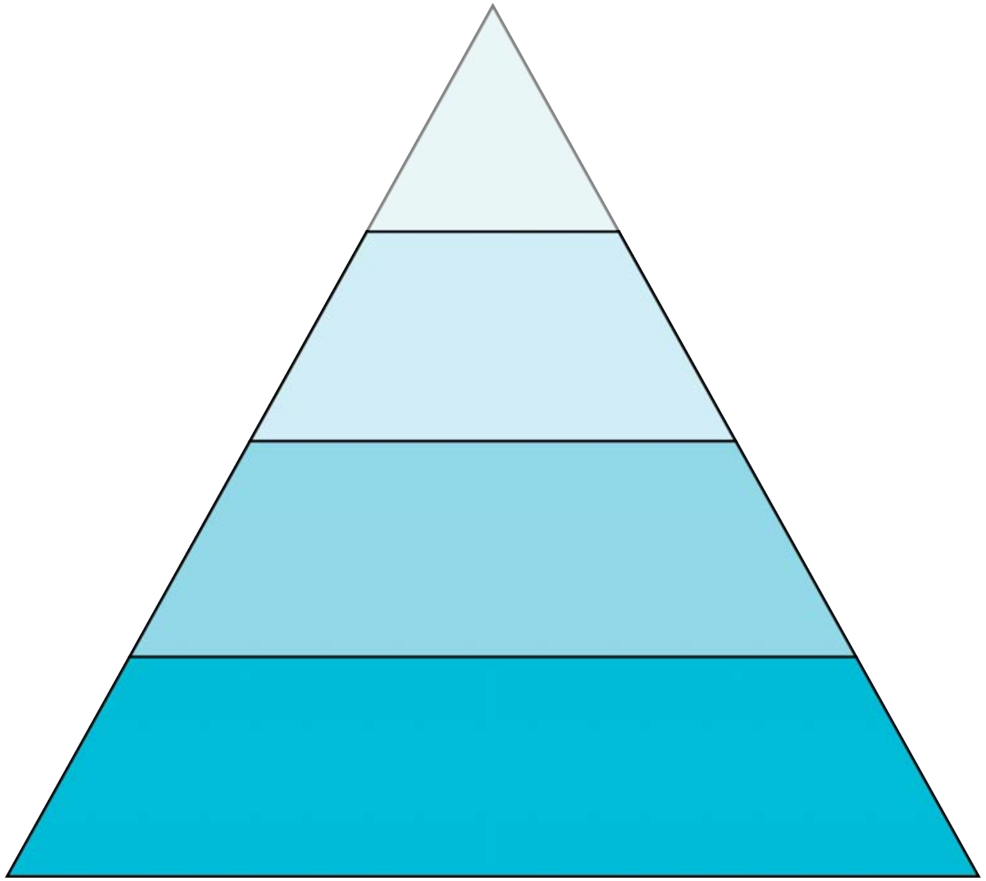
Model
(CESM 1°)

Observational Estimate
(Landschützer et al.)



Long et al., *JAMES* (2021)

Modeling frameworks to support ocean CDR & MRV



CDR processes & experimental framework

Observing system design
Verification with sequential data assimilation

Data assimilation

Ocean physical and biogeochemical models

Data assimilation

Free-running models display systematic bias

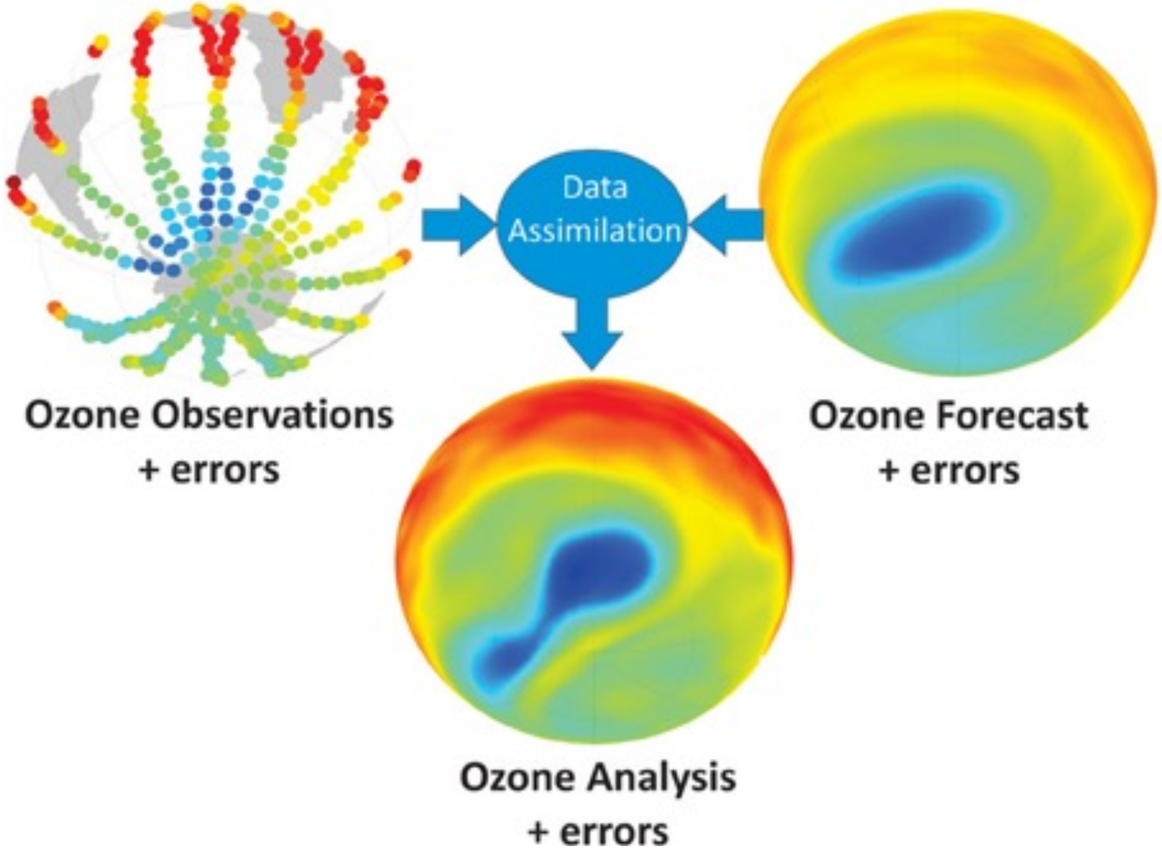
DA can produce a model-data state estimate ("reanalysis"), reducing bias

DA provides model states that are maximally consistent with observations

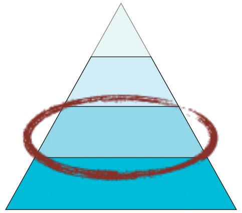
DA can provide explicit uncertainty estimates and assessments of which observations are most important

Two approaches:

- Sequential ensemble DA (filter)
- Variational methods (smoother)

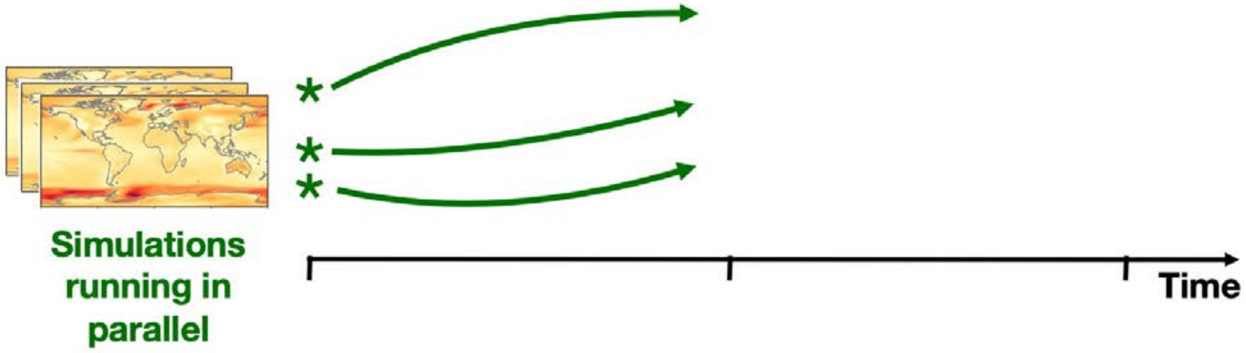


Lahoz & Schneider (2014)



Ensemble data assimilation

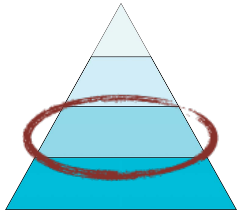
Integrate ensemble of models



Courtesy of D. Amrhein (NCAR)

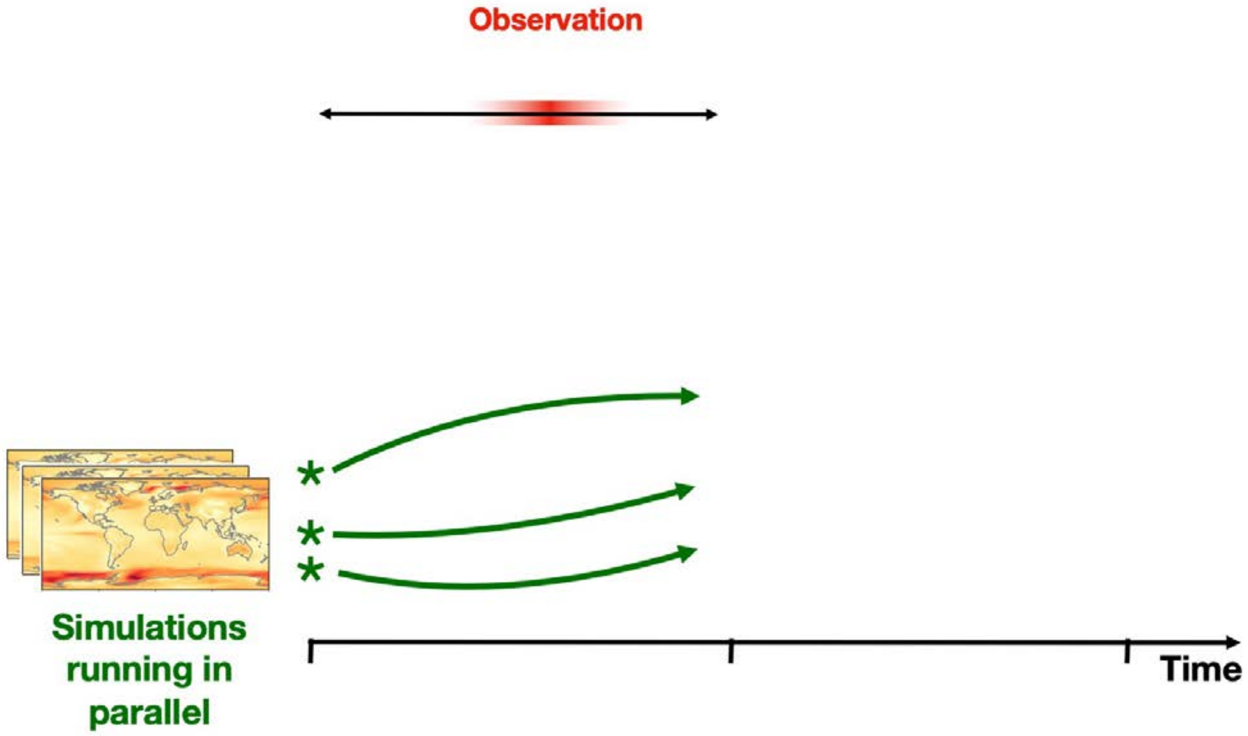
Anderson, et al. (2007)

Ensemble sequential data assimilation



Ensemble data assimilation

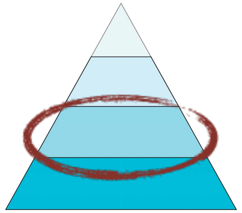
Integrate ensemble of models



Courtesy of D. Amrhein (NCAR)

Anderson, et al. (2007)

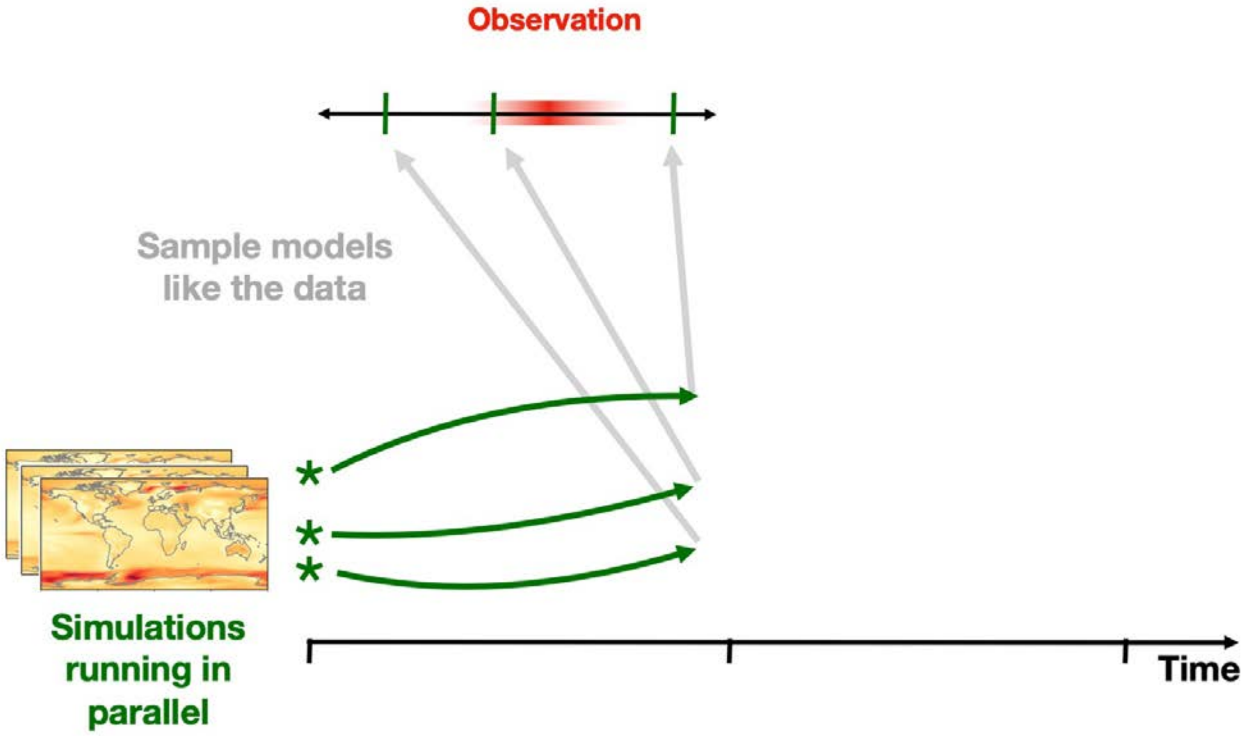
Ensemble sequential data assimilation



Ensemble data assimilation

Integrate ensemble of models

Sample models like the data



Courtesy of D. Amrhein (NCAR)

Anderson, et al. (2007)

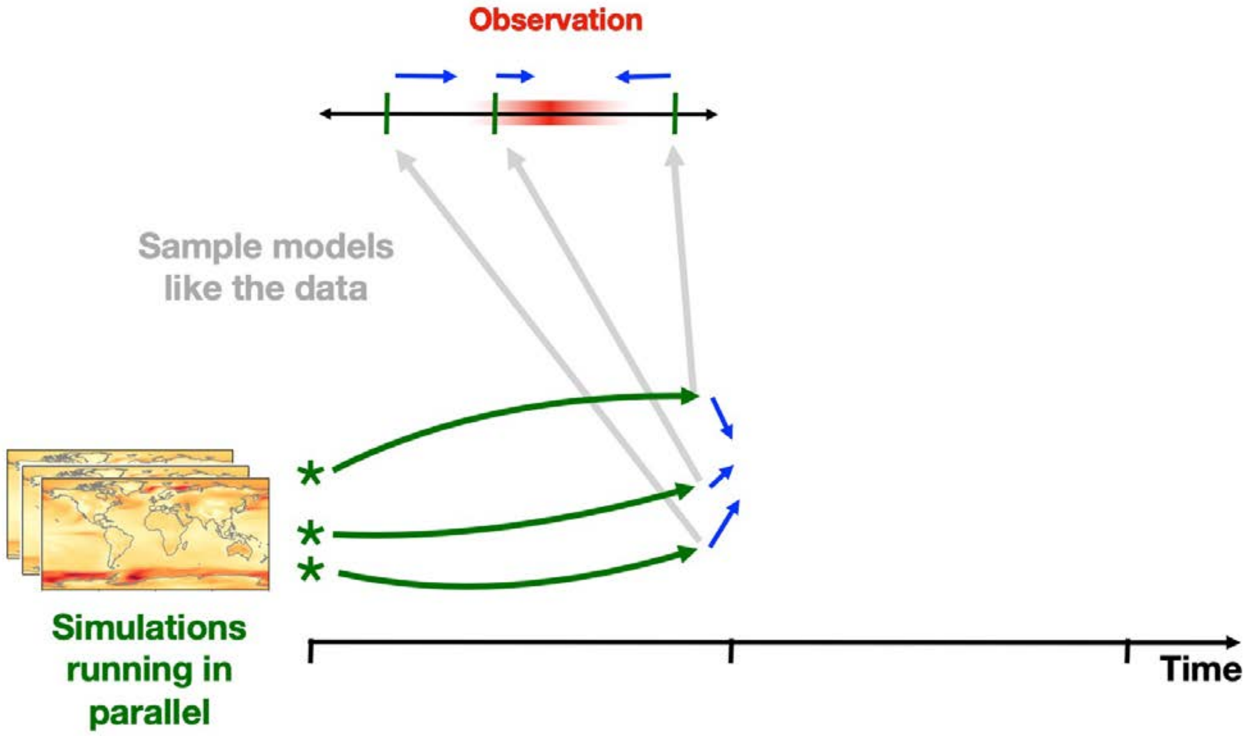
Ensemble sequential data assimilation

Ensemble data assimilation

Integrate ensemble of models

Sample models like the data

Add “increments” to model states to improve fit to data



Courtesy of D. Amrhein (NCAR)

Anderson, et al. (2007)

Ensemble sequential data assimilation

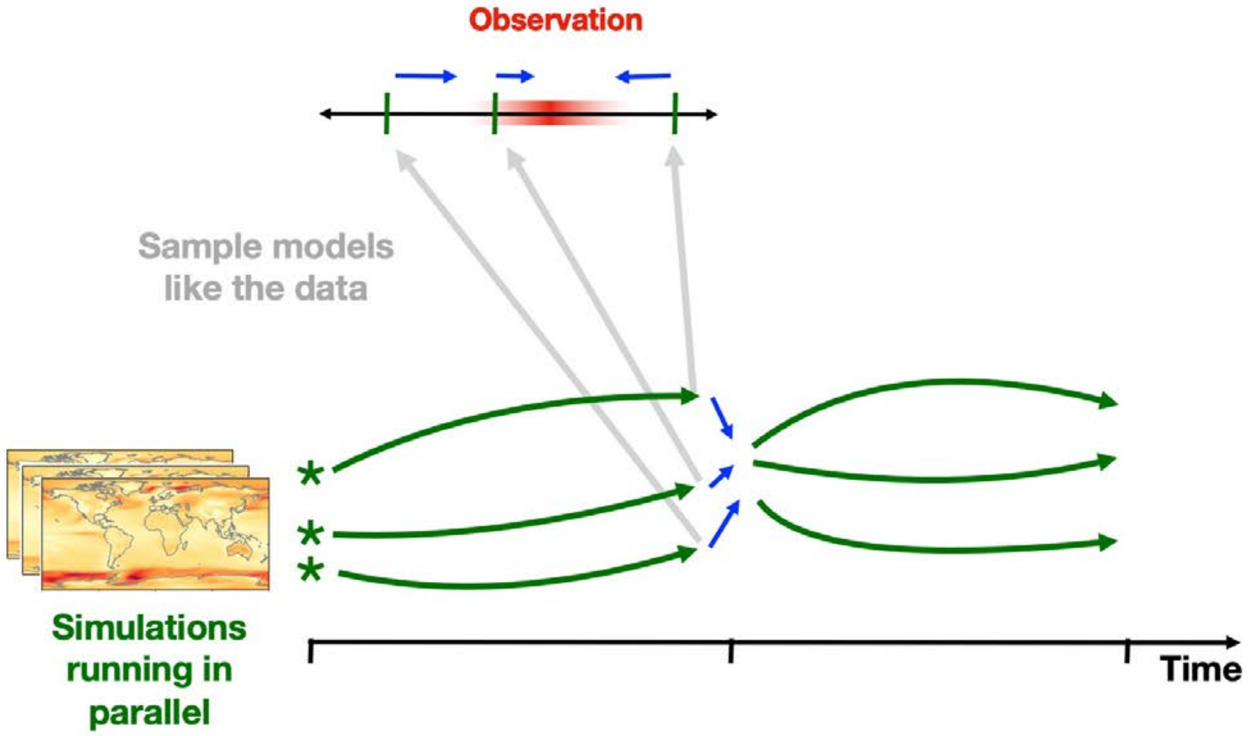
Ensemble data assimilation

Integrate ensemble of models

Sample models like the data

Add “increments” to model states to improve fit to data

Continue integration



Courtesy of D. Amrhein (NCAR)

Anderson, et al. (2007)

Ensemble sequential data assimilation

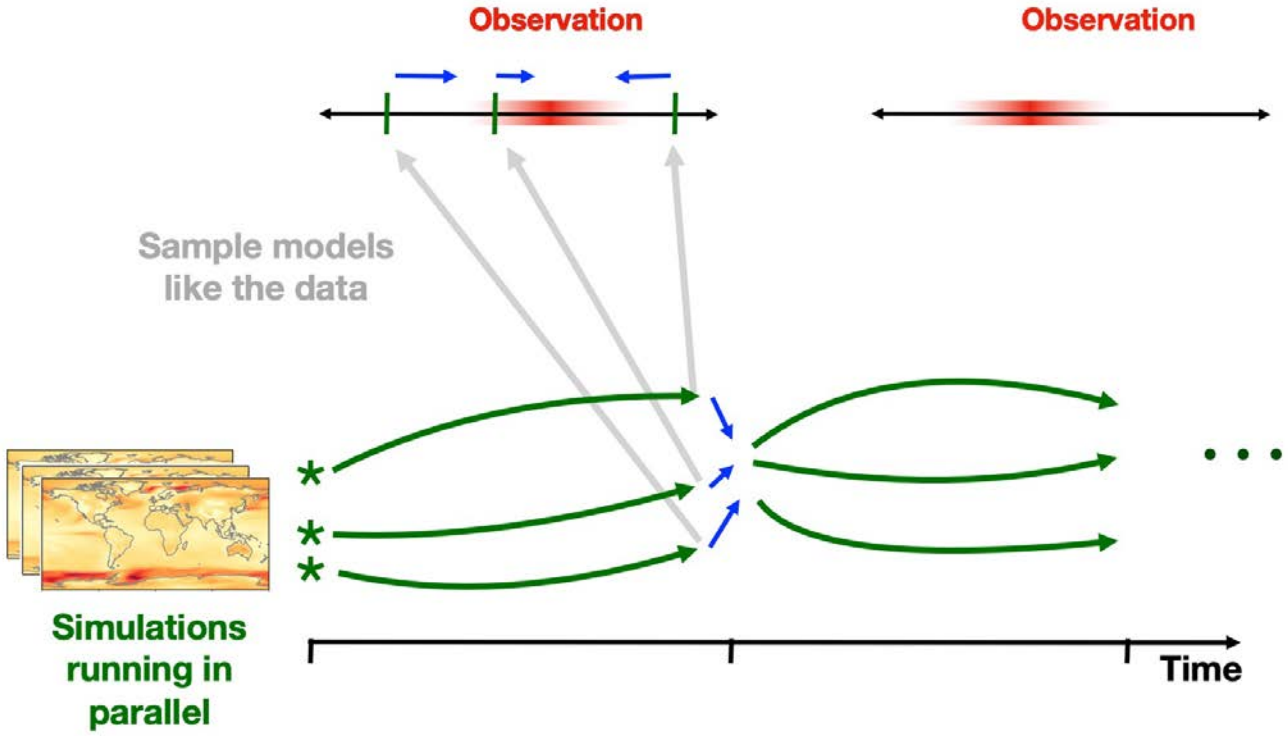
Ensemble data assimilation

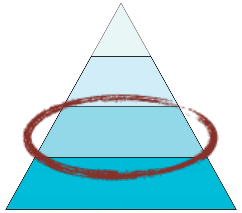
Integrate ensemble of models

Sample models like the data

Add "increments" to model states to improve fit to data

Continue integration

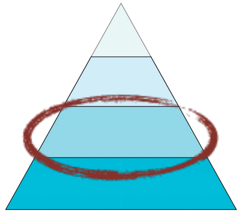




Courtesy of D. Amrhein (NCAR)

Anderson, et al. (2007)

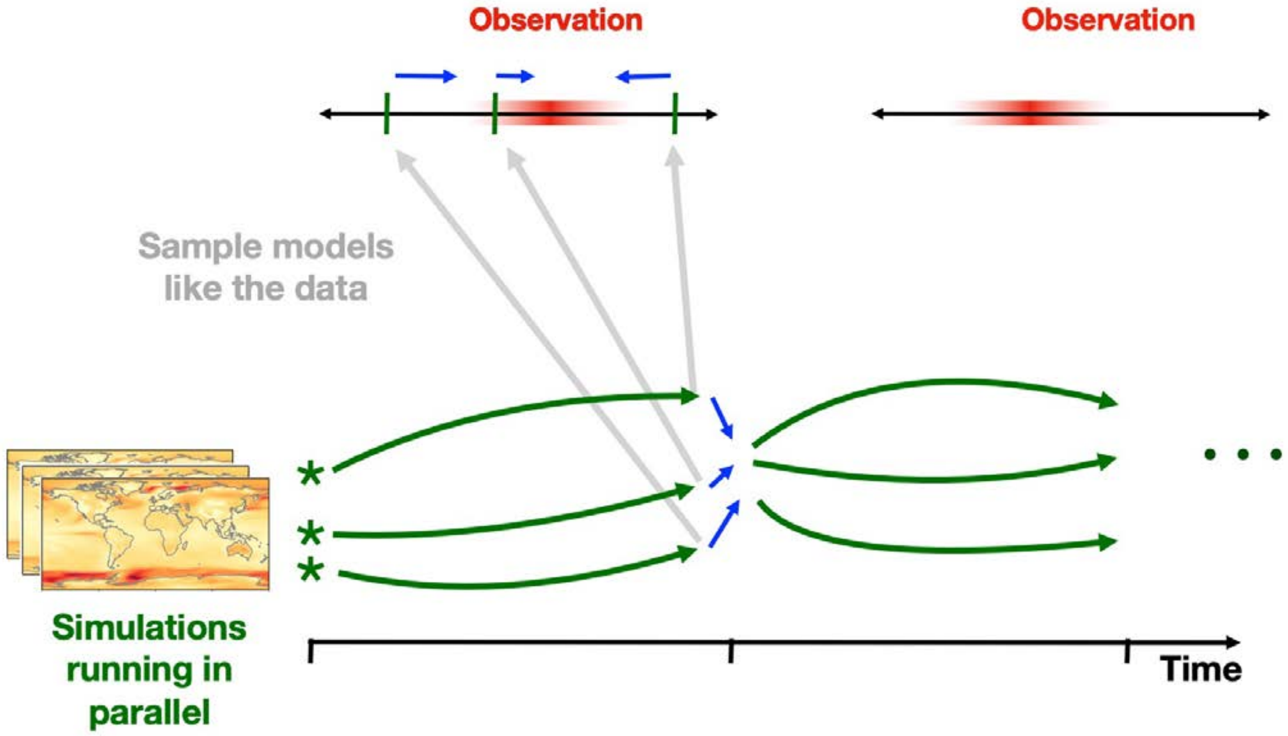
Ensemble sequential data assimilation



Challenges and shortcomings

Adding increments to the state can break the model physics

Two often-necessary tools: inflation and localization



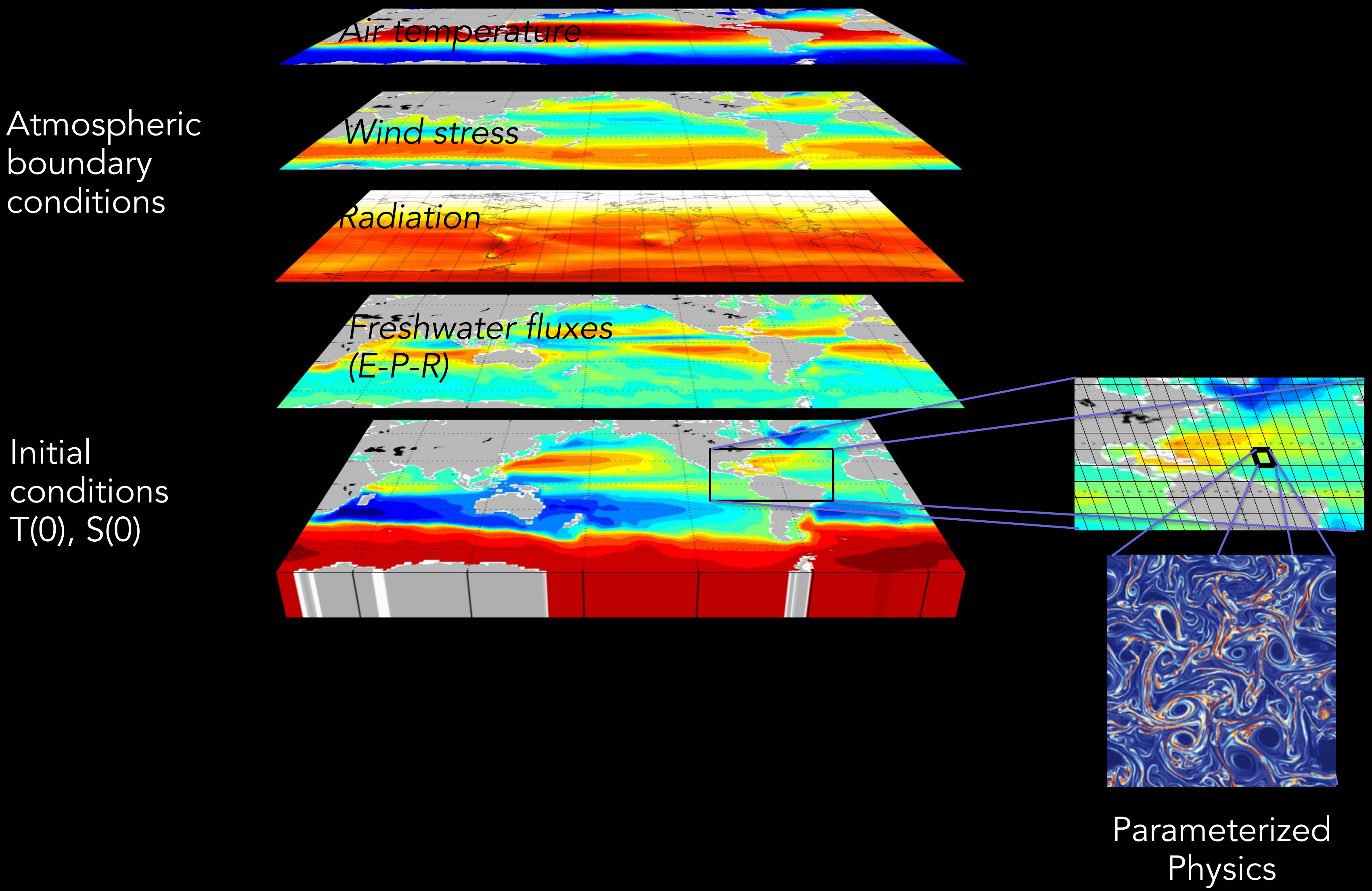
Courtesy of D. Amrhein (NCAR)

Anderson, et al. (2007)

Variational methods solve an iterative optimization problem to minimize an objective function

Example: the ECCO state estimate

Control variables, **u**



Variational methods in practice: the ECCO state estimate

Inverse problem solved by ECCO

Solve for a set of

- initial conditions,
- atmospheric boundary forcing,
- ocean mixing parameters,

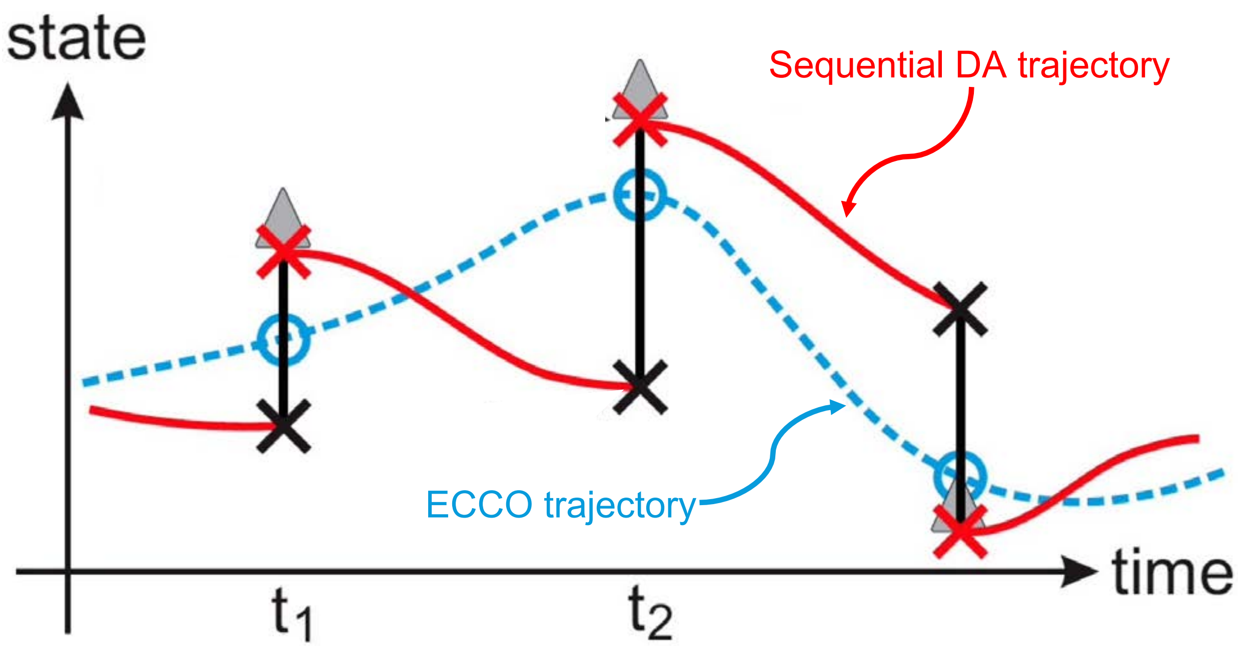
to minimize residuals between the model solution and the observations in a least-squares sense.

Advantages

- Conserves ocean properties
- Uses information from past and future to constrain state (smoother rather than a filter)

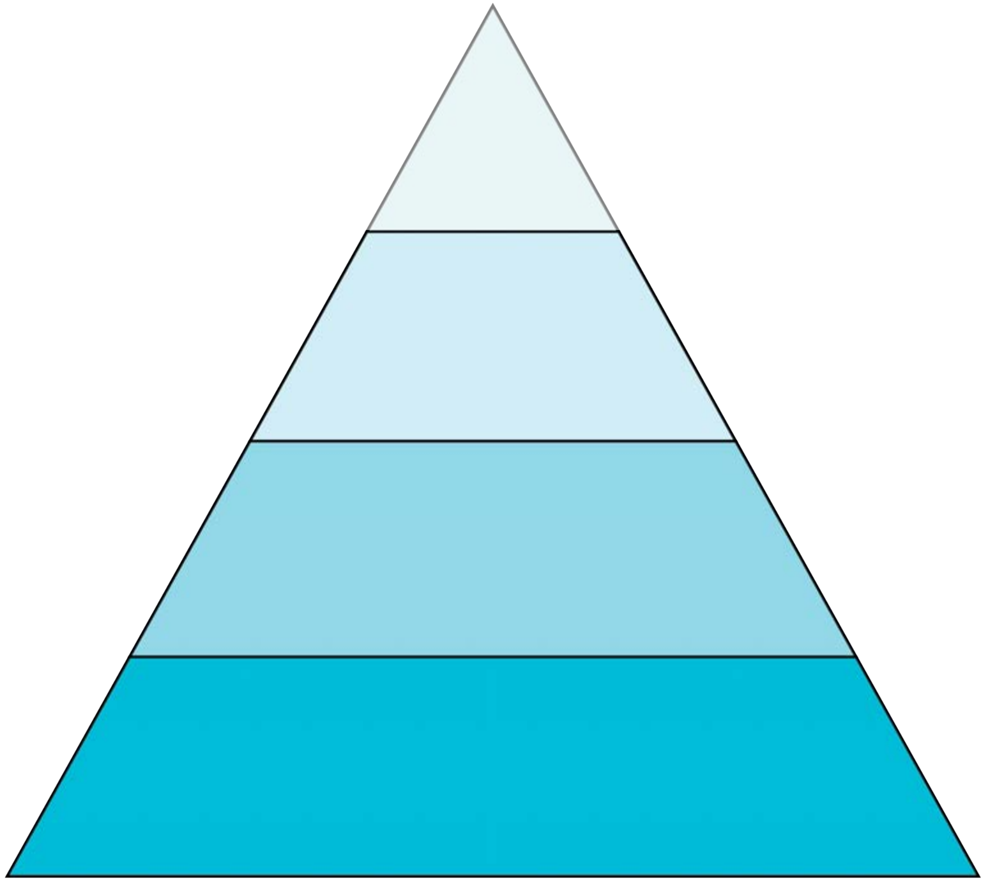
Disadvantages

- Can generate spurious fluxes, e.g. in Lab. Sea
- No “errors of the day”
- No uncertainty quantification
- Expensive to maintain adjoint



Courtesy of D. Amrhein (NCAR) & D. Menemenlis (JPL)

Modeling frameworks to support ocean CDR & MRV



CDR processes & experimental framework

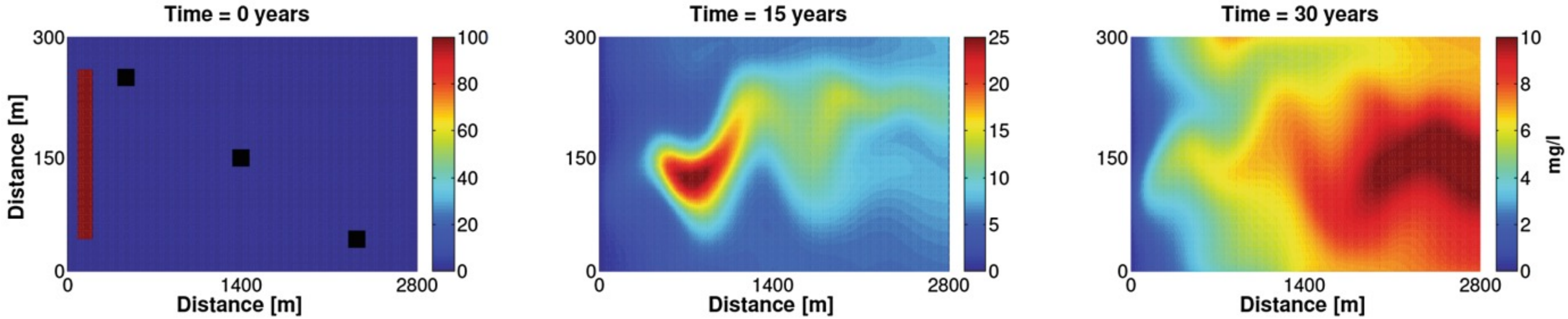
Observing system design
Verification with sequential data assimilation

Data assimilation

Ocean physical and biogeochemical models

Optimization of observing location

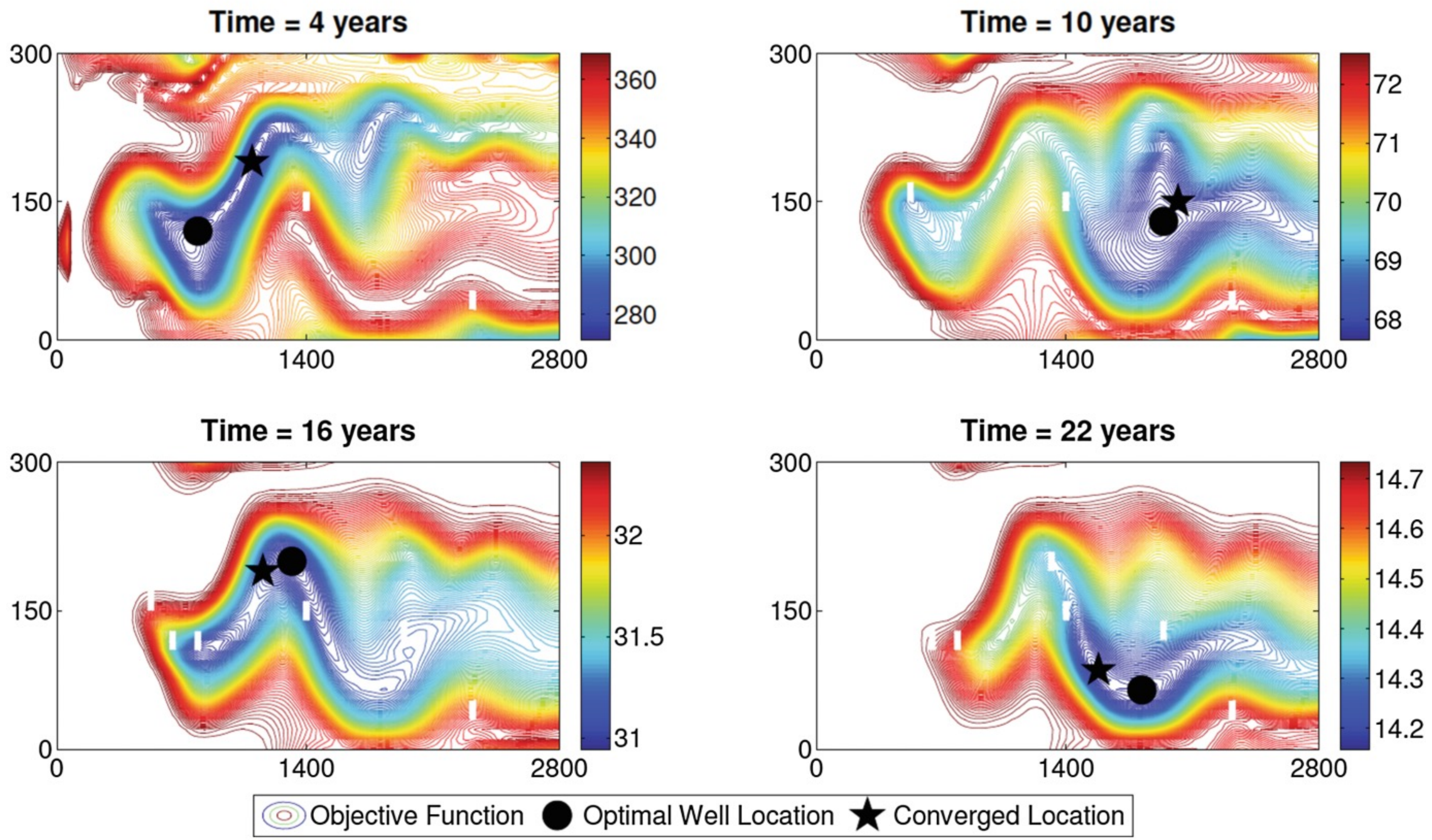
What is an optimal sampling strategy to best constrain a quantity of interest (i.e., “verification variable”)?



Gharamti et al. (2014)

Optimization of observing location

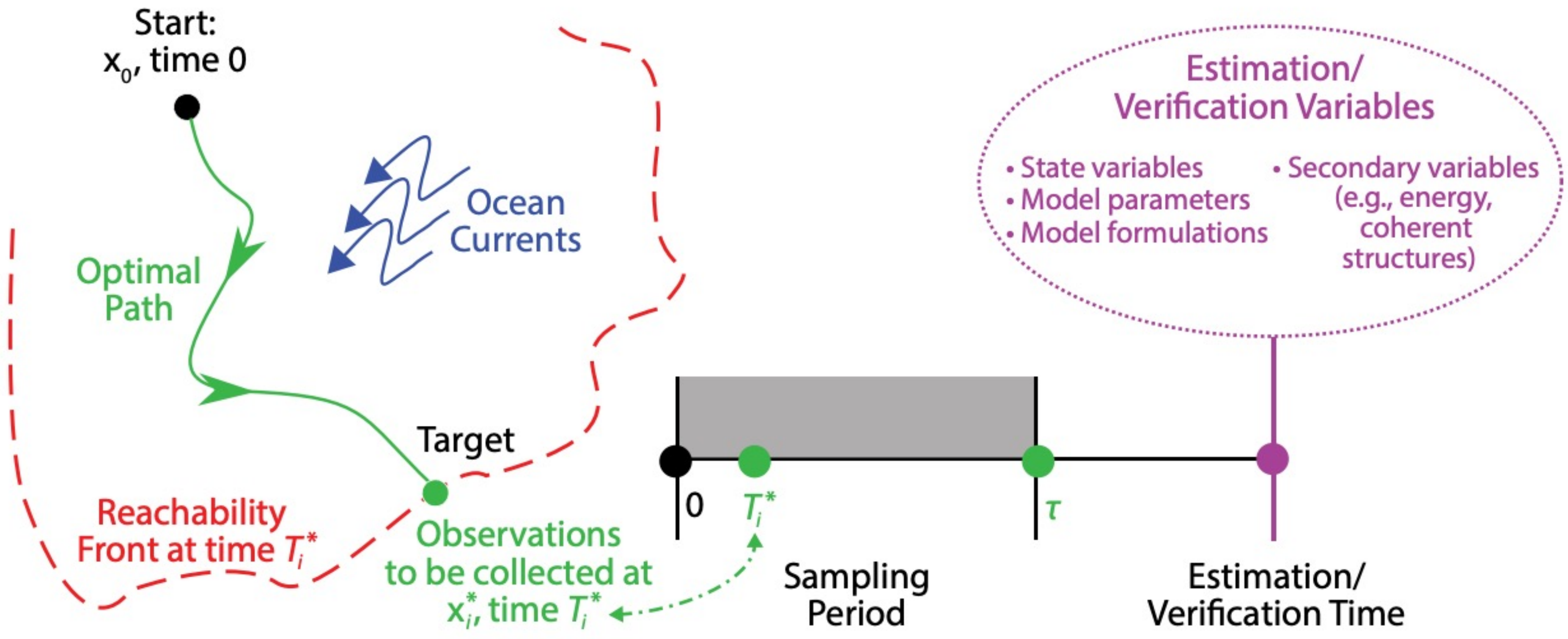
What is an optimal sampling strategy to best constrain a quantity of interest (i.e., “verification variable”)?



Gharamti et al. (2014)

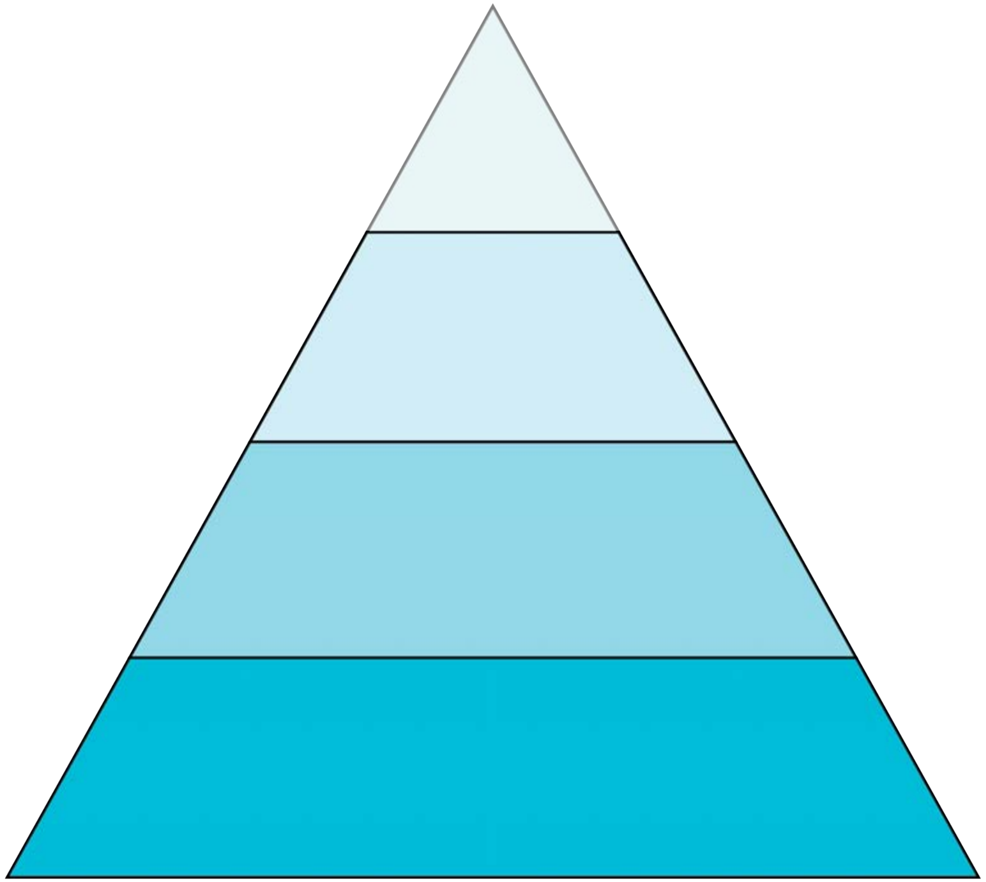
Optimal observing system design

What is an optimal sampling strategy to best constrain a quantity of interest (i.e., “verification variable”)?



Lermusiaux, et al. (2017)

Modeling frameworks to support ocean CDR & MRV



CDR processes & experimental framework

Observing system design
Verification with sequential data assimilation

Data assimilation

Ocean physical and biogeochemical models

MRV Challenges

Unfavorable signal-to-noise ratios

- Large background fields
- Dynamic variability

Establishing baseline counterfactual to assess additionality

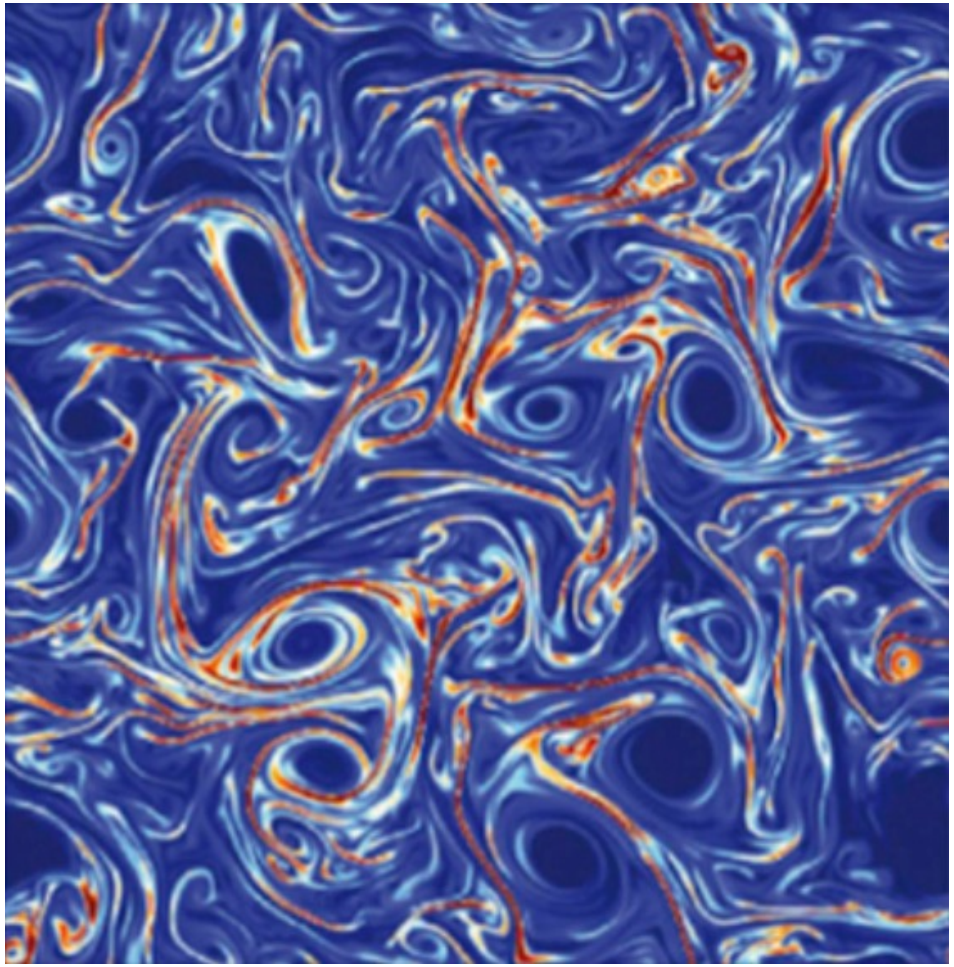
Large spatiotemporal scales

- Slow CO_2 equilibration timescale
- Large-scale overturning circulation

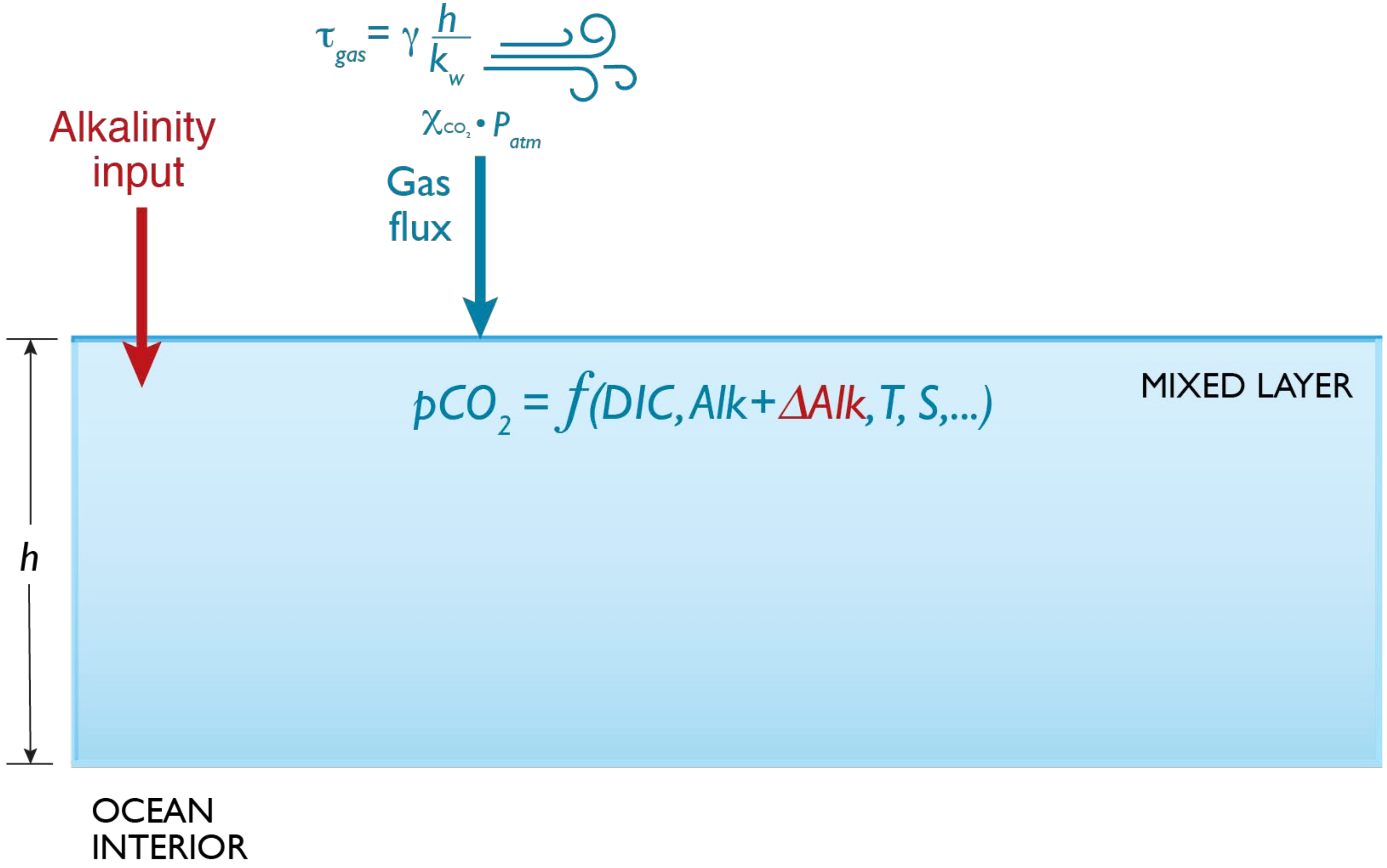
Model skill and parameter uncertainty

Unknown unknowns

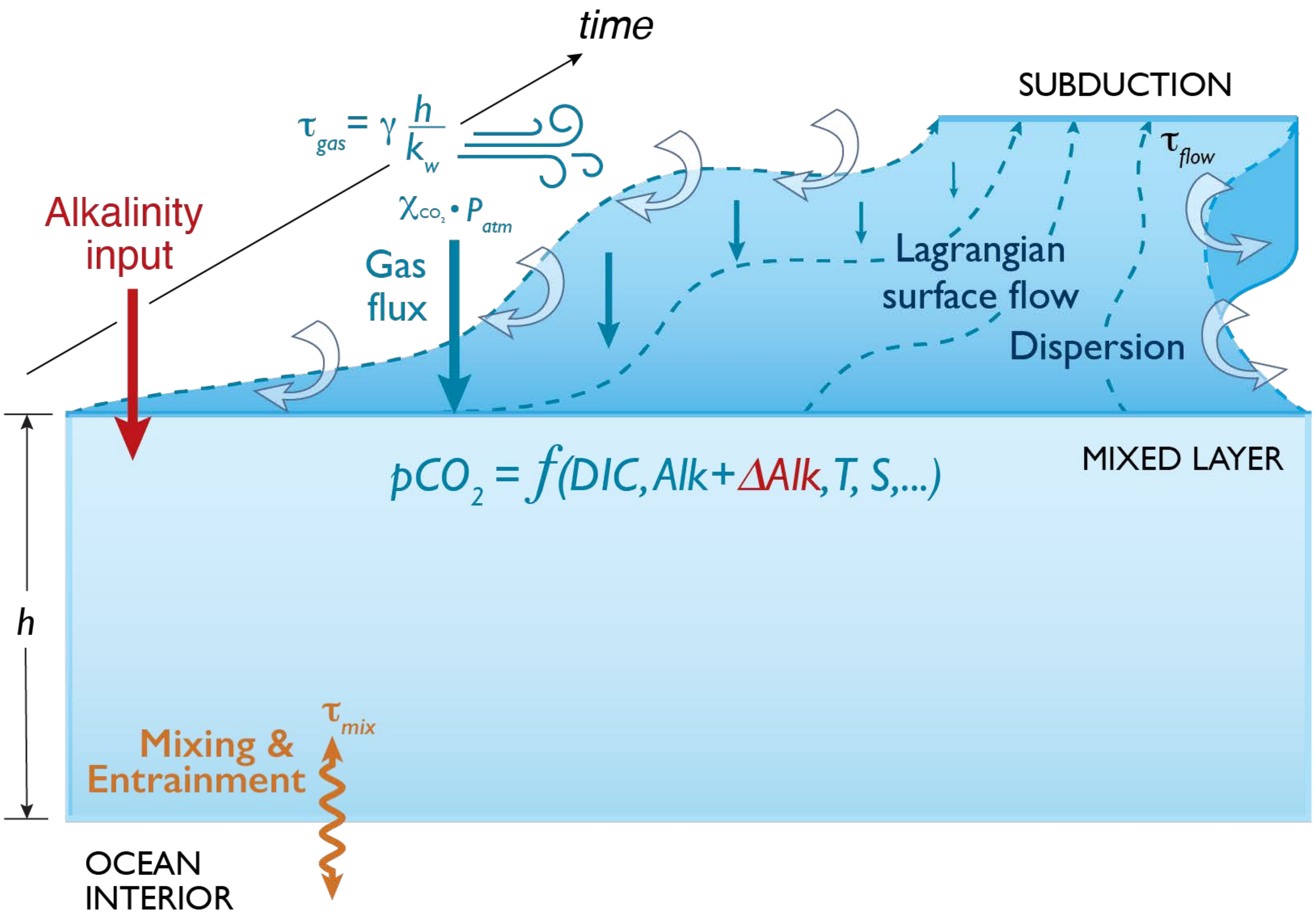
- Ecosystem compensation



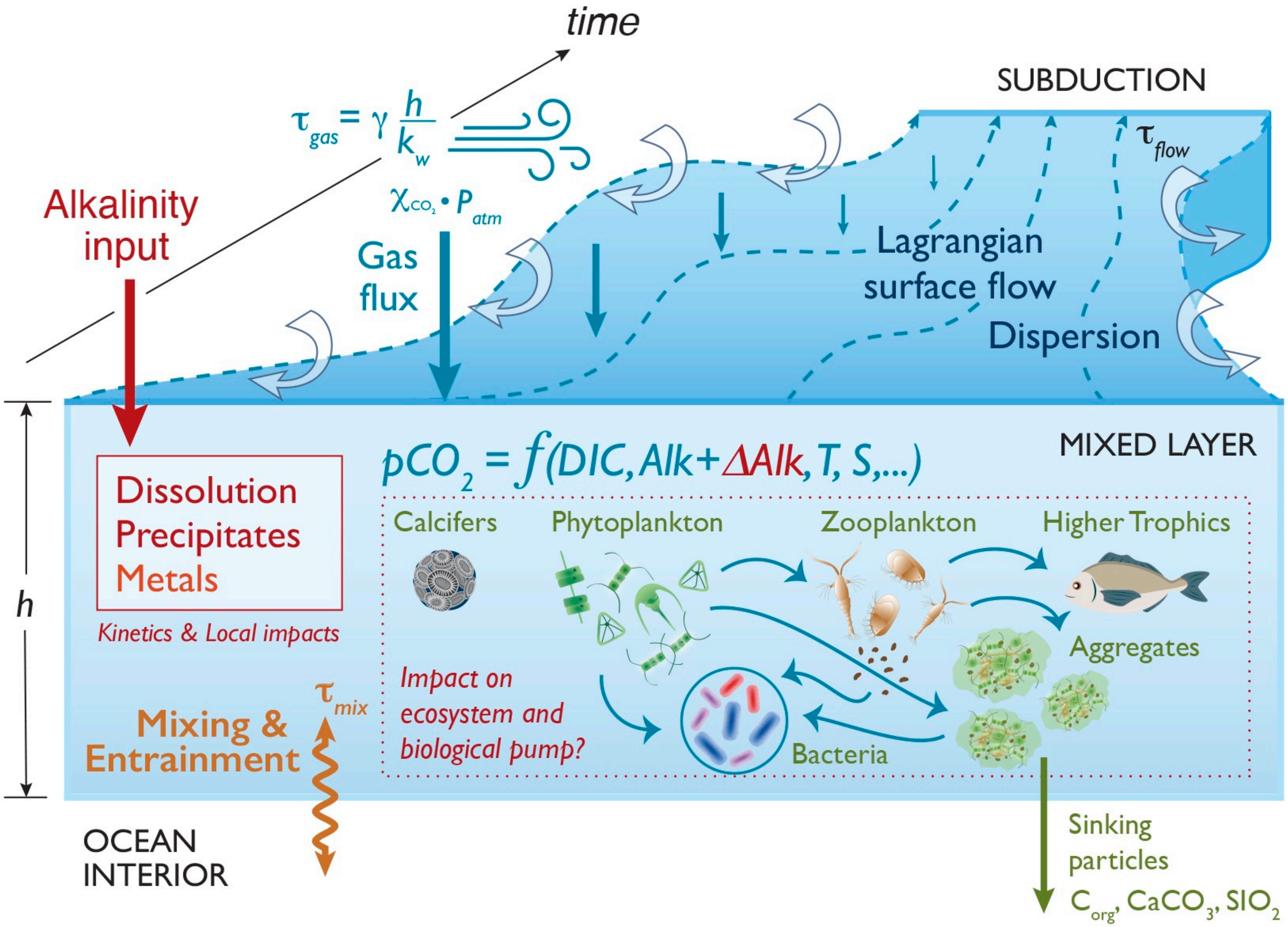
Ocean alkalinity enhancement



Ocean alkalinity enhancement

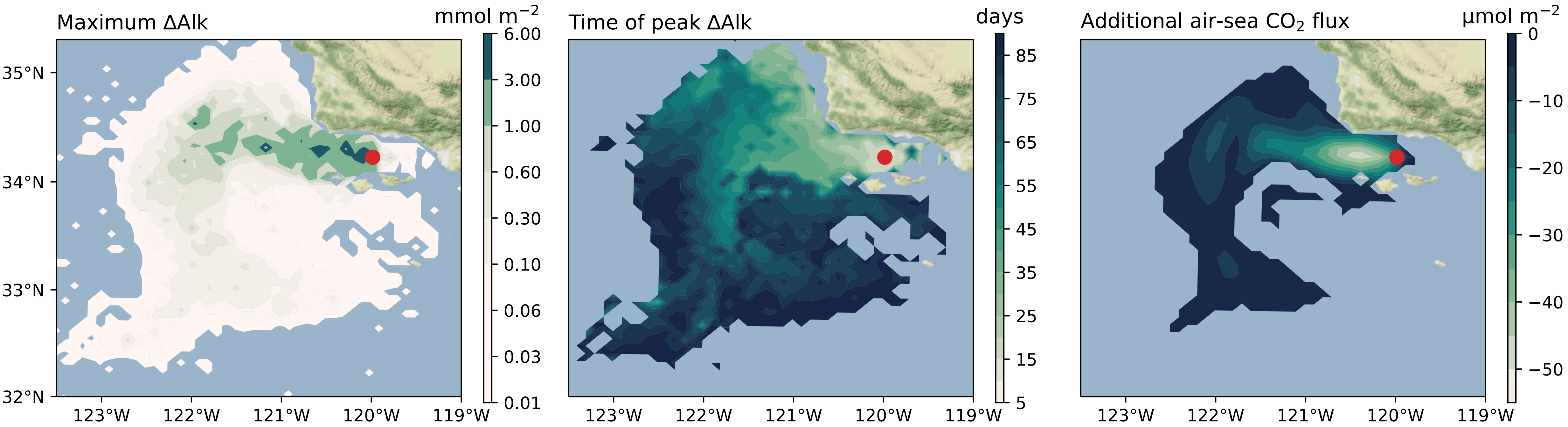


Ocean alkalinity enhancement



Results from idealized experiments

Simulated alkalinity release (CESM 0.1°)



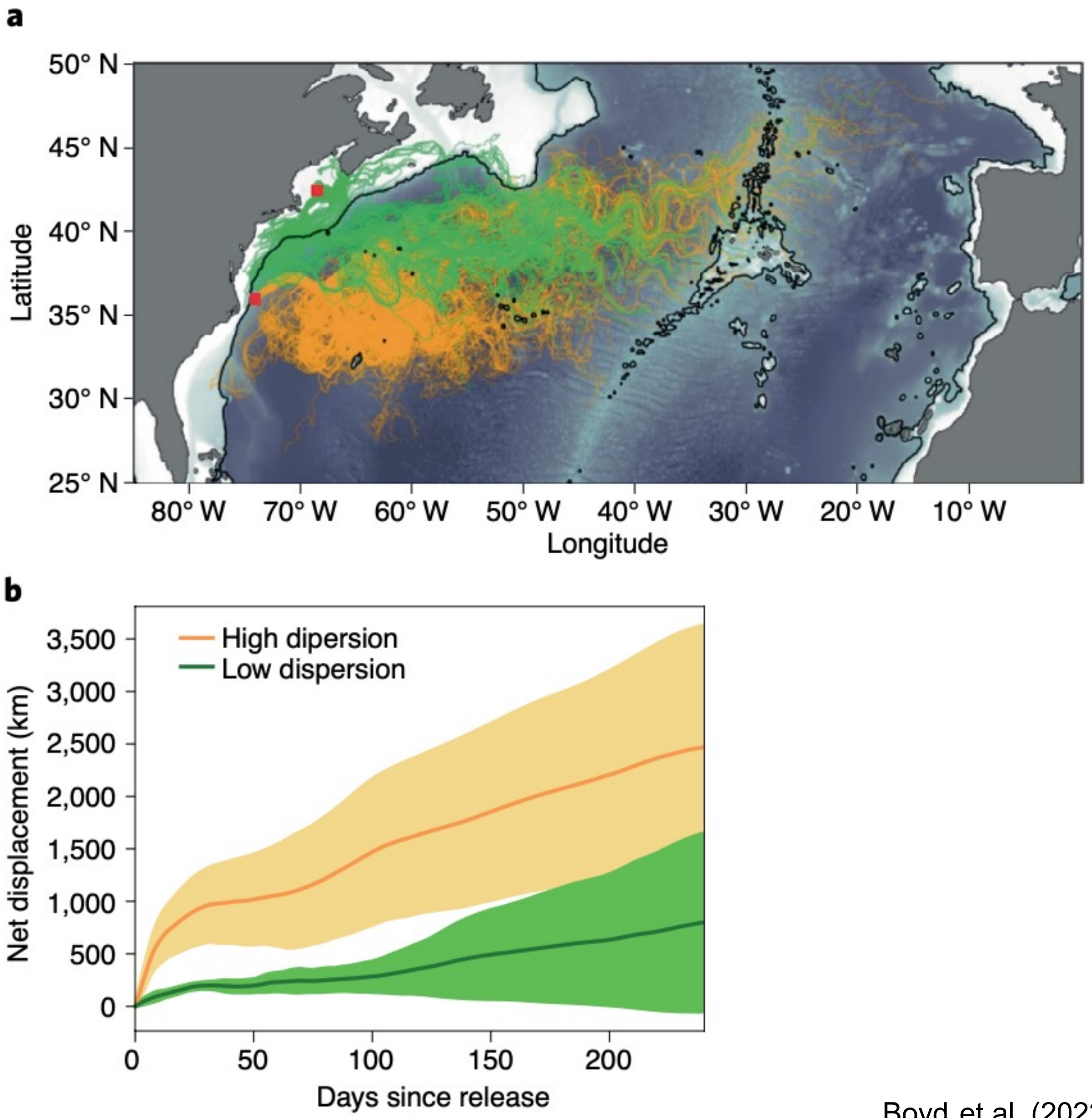
Particle dispersion modeling

Lagrangian (i.e., flow-following) frameworks provide complementary information to GCMs

Particles can be advected in pre-computed velocity fields

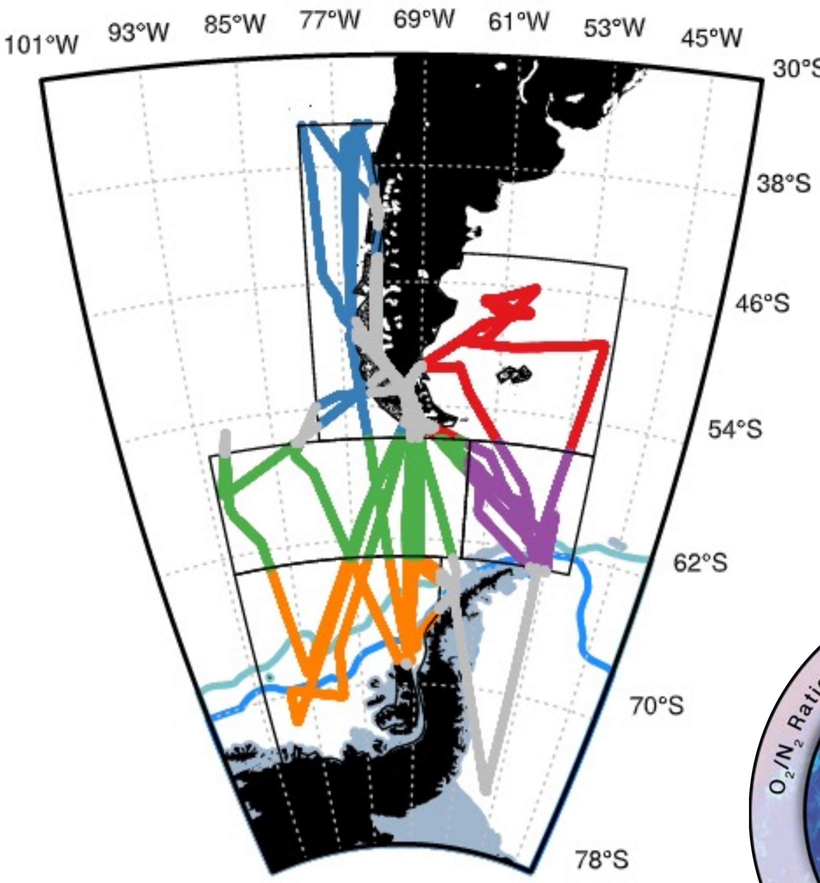
Quantify dispersion and trajectories

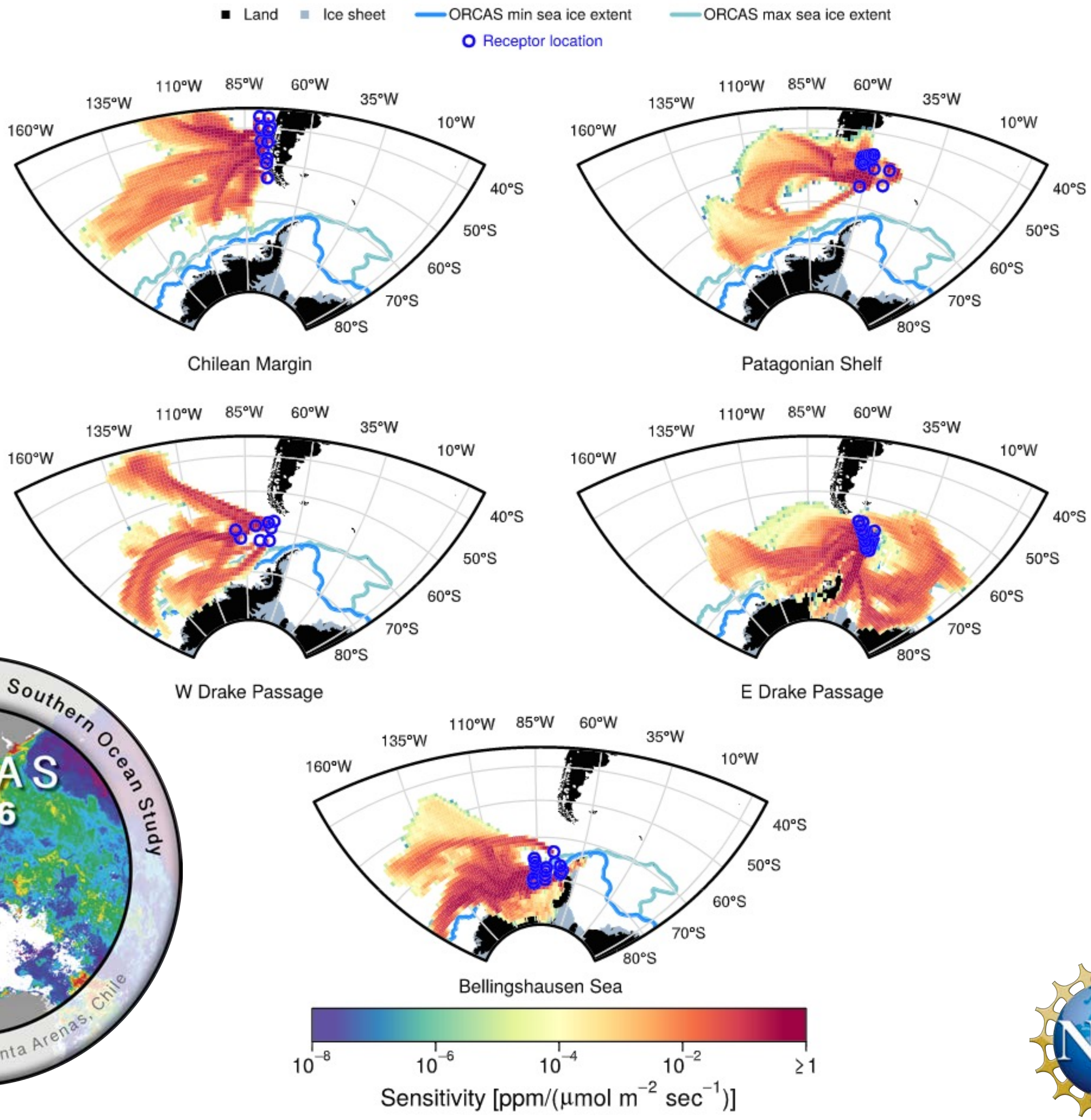
Possible to sample properties and compute transformations along particle paths



Boyd et al. (2022)

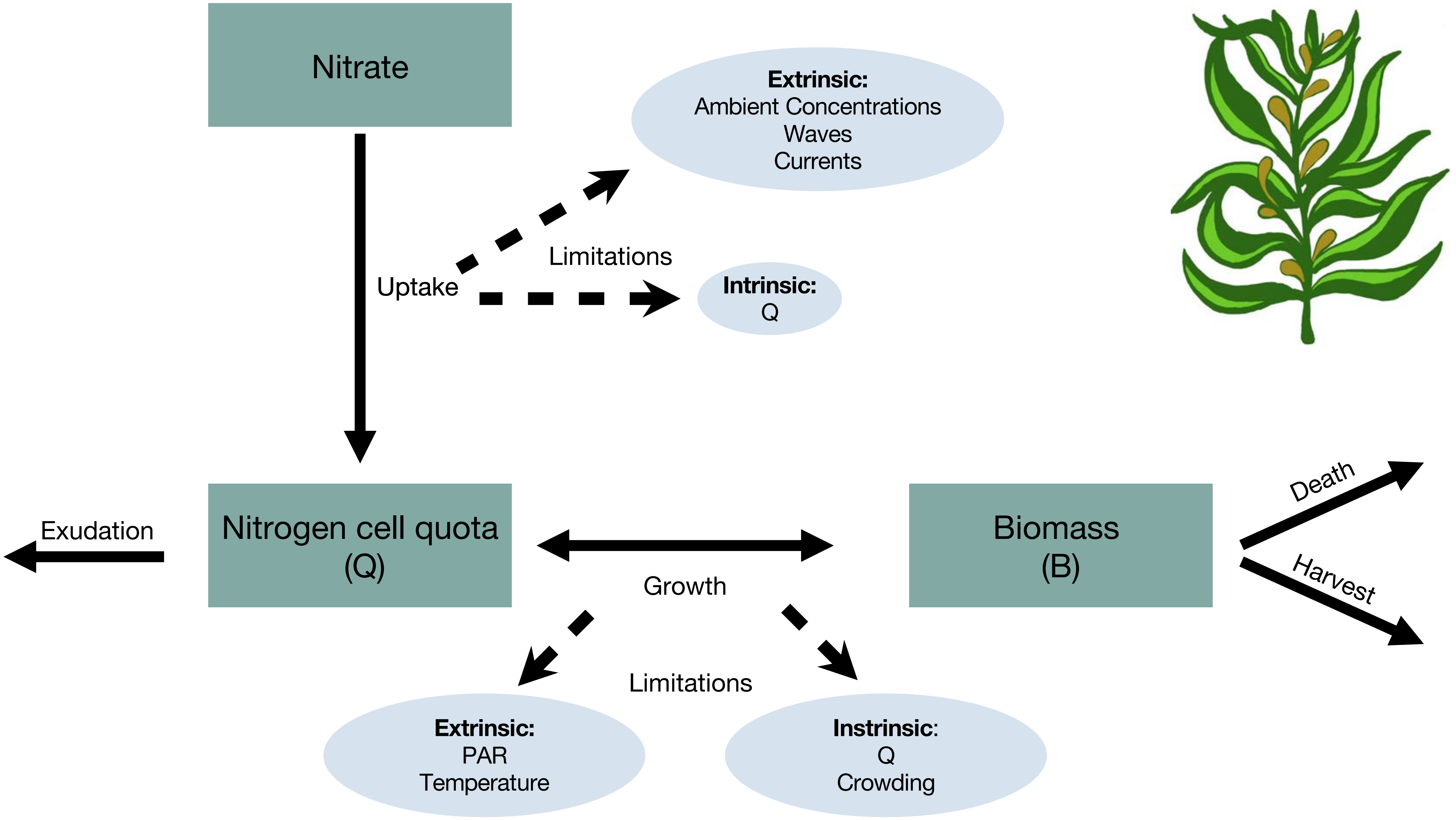
Lagrangian back-trajectories & footprints computed for aircraft samples





Morgan et al., *JGR*, (2019)

Global MacroAlgae Cultivation MODelling System (G-MACMODS)



Courtesy of I. Arzeno (UCI)

Frieder et al. (2022)

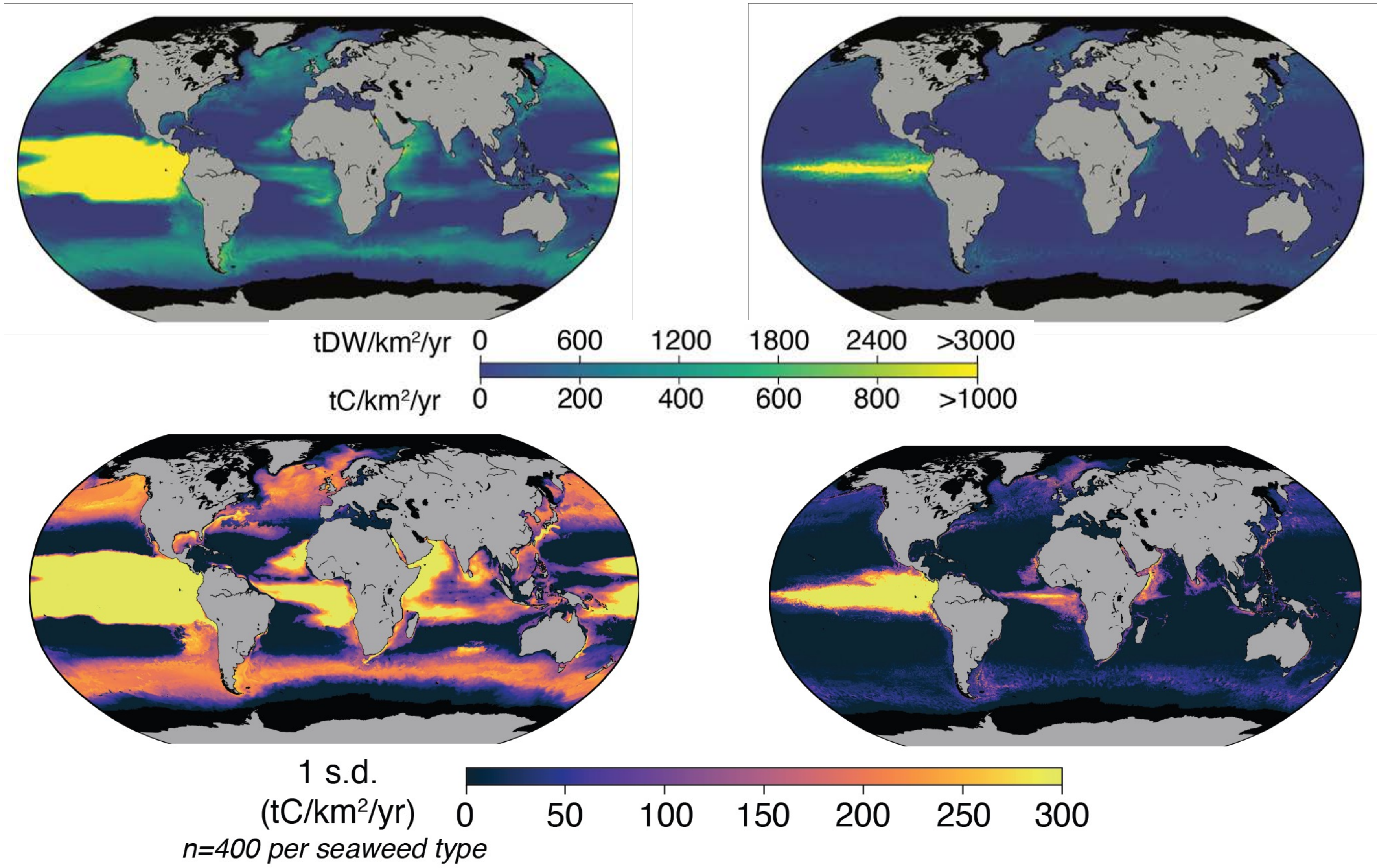
Projecting potential macroalgae yields: uncertainty in parameters

Uncertainty analysis around biological parameters:

- Maximum uptake rate
- Half saturation constant
- Compensation irradiance
- Saturation irradiance
- Maximum growth rate
- Crowding
- Ratio of biomass to surface area
- Minimum nitrogen cell quota
- Max nitrogen cell quota
- Drag coefficient
- Nitrogen exudation rate
- Mortality rate

n = 400 Monte Carlo runs for each seaweed group and each nutrient scenario (ambient and flux-limited)

Ambient nutrients | **Flux-limited nutrients**

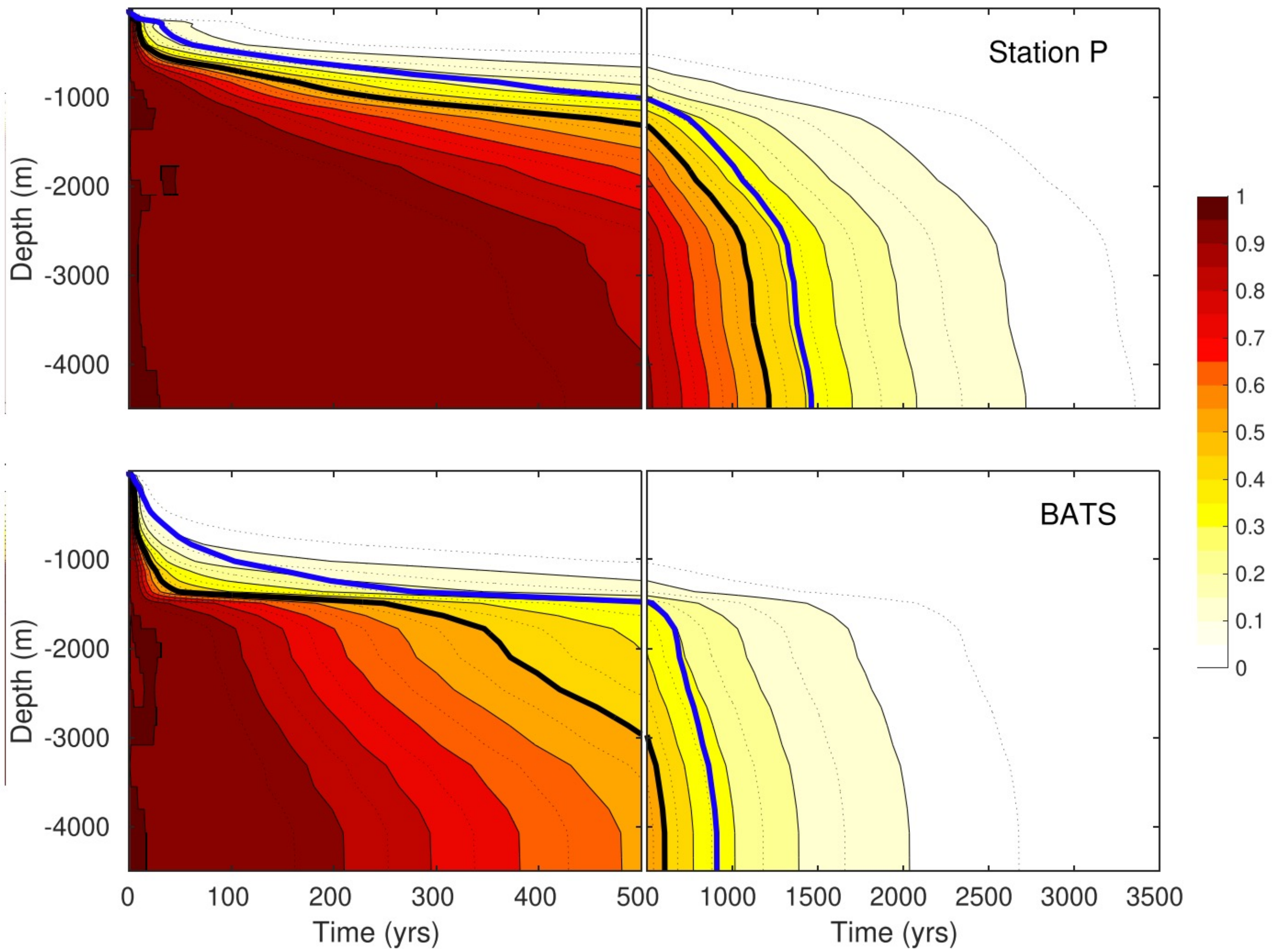


Courtesy of I. Arzeno (UCI)

Arzeno et al. (in revision)

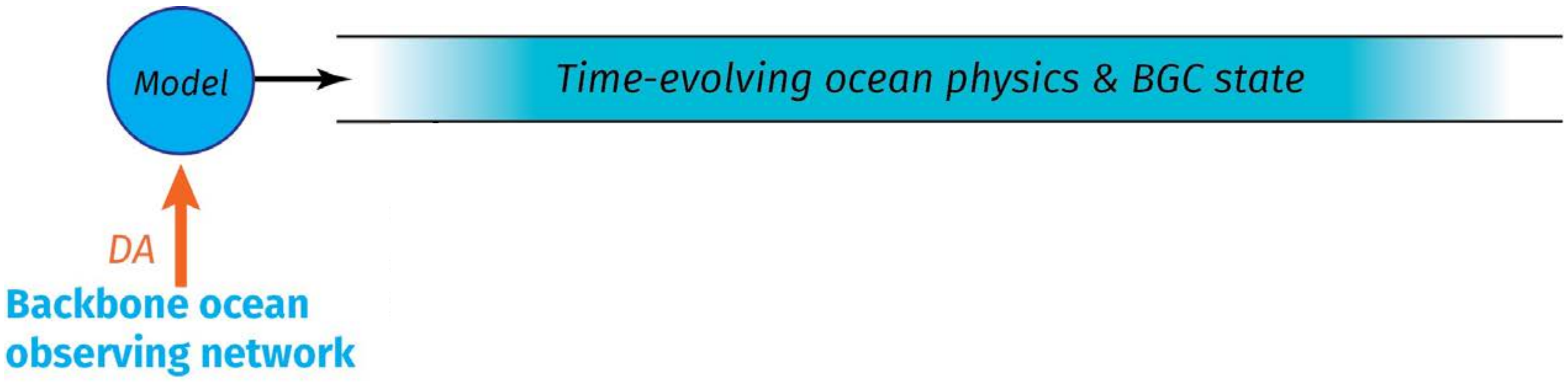
Separation of concerns: mapping re-emergence timescales

Fraction of carbon remaining sequestered after injection at depth

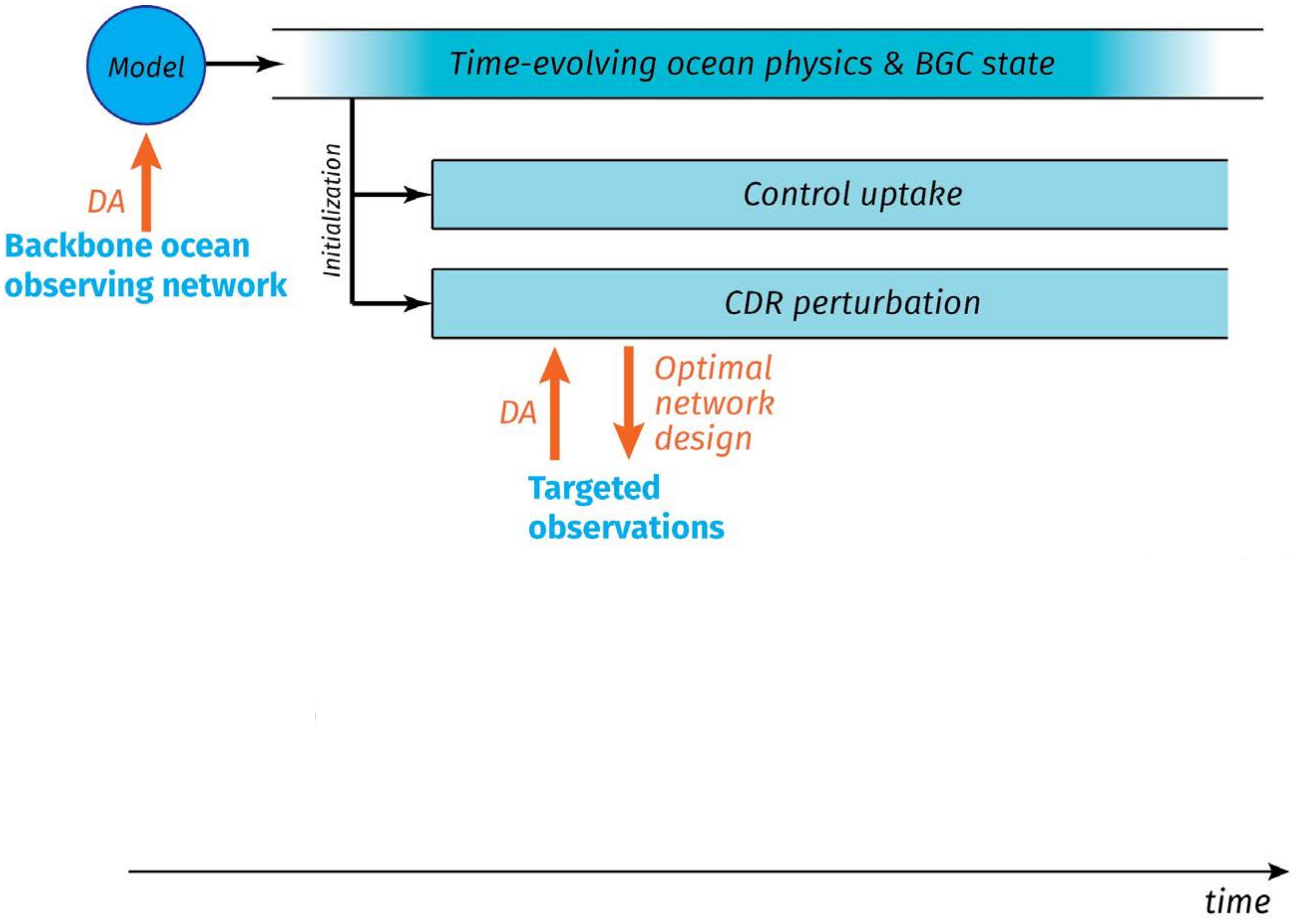


Siegel et al. (2022)

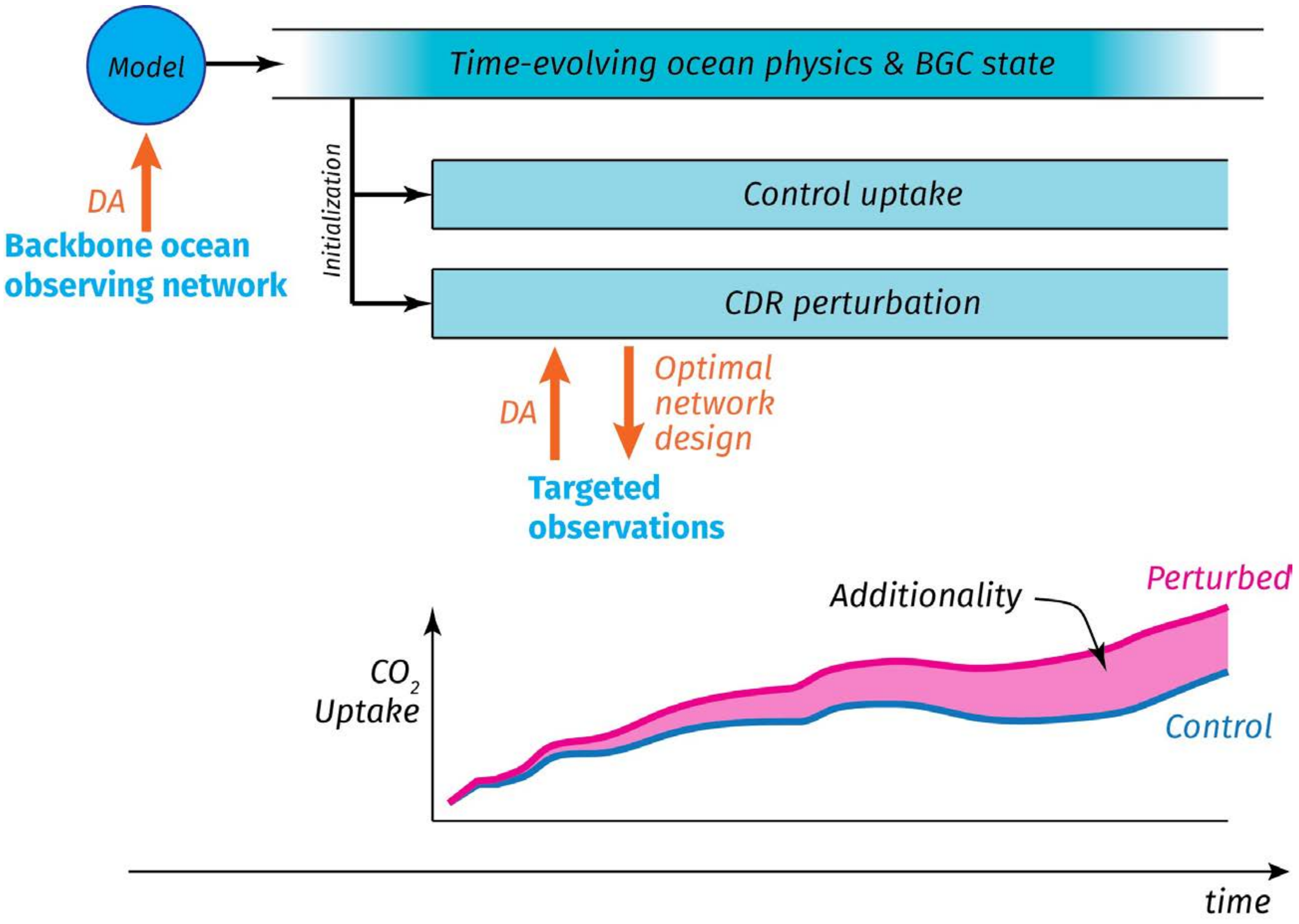
Experimental framework for MRV



Experimental framework for MRV



Experimental framework for MRV



Computation and data constraints

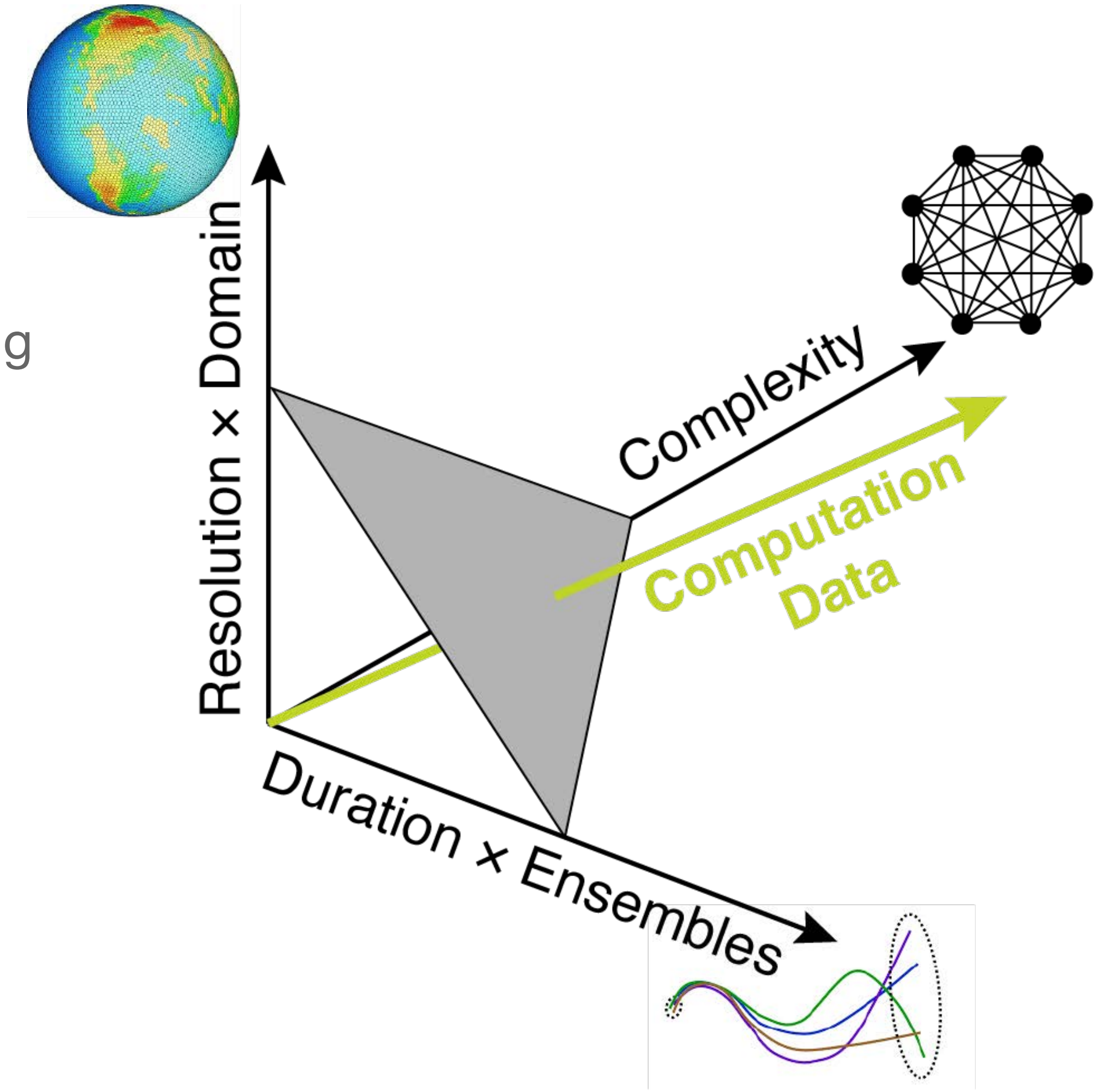
Computational constraints

Develop creative solutions

- E.g., geophysical modeling + machine learning

Need hierarchy of modeling tools tailored to questions of interest

- Modeling systems built from interoperable components
- Enable bespoke sandboxes



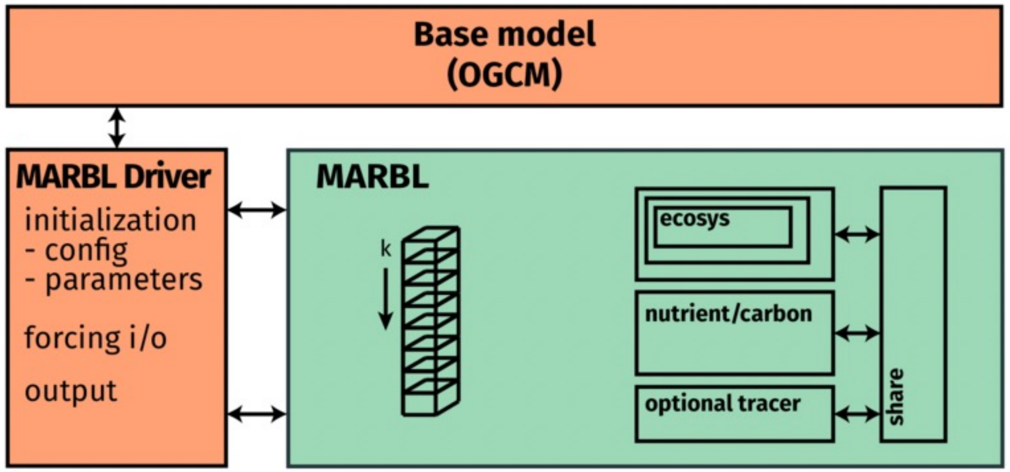
Interoperability between modular components

Modularity: facilitated by coupling infrastructure

Enable “mix-and-match” capabilities

Leverage interoperability to
- Isolate and study processes
- Explore structural uncertainty
- Leverage communities of practice

Marine Biogeochemical Library (MARBL)



$$\frac{\partial \varphi}{\partial t} + \nabla \cdot (\vec{u}\varphi) - \nabla \cdot (K \nabla \varphi) = J(\varphi)$$

Computation and data constraints

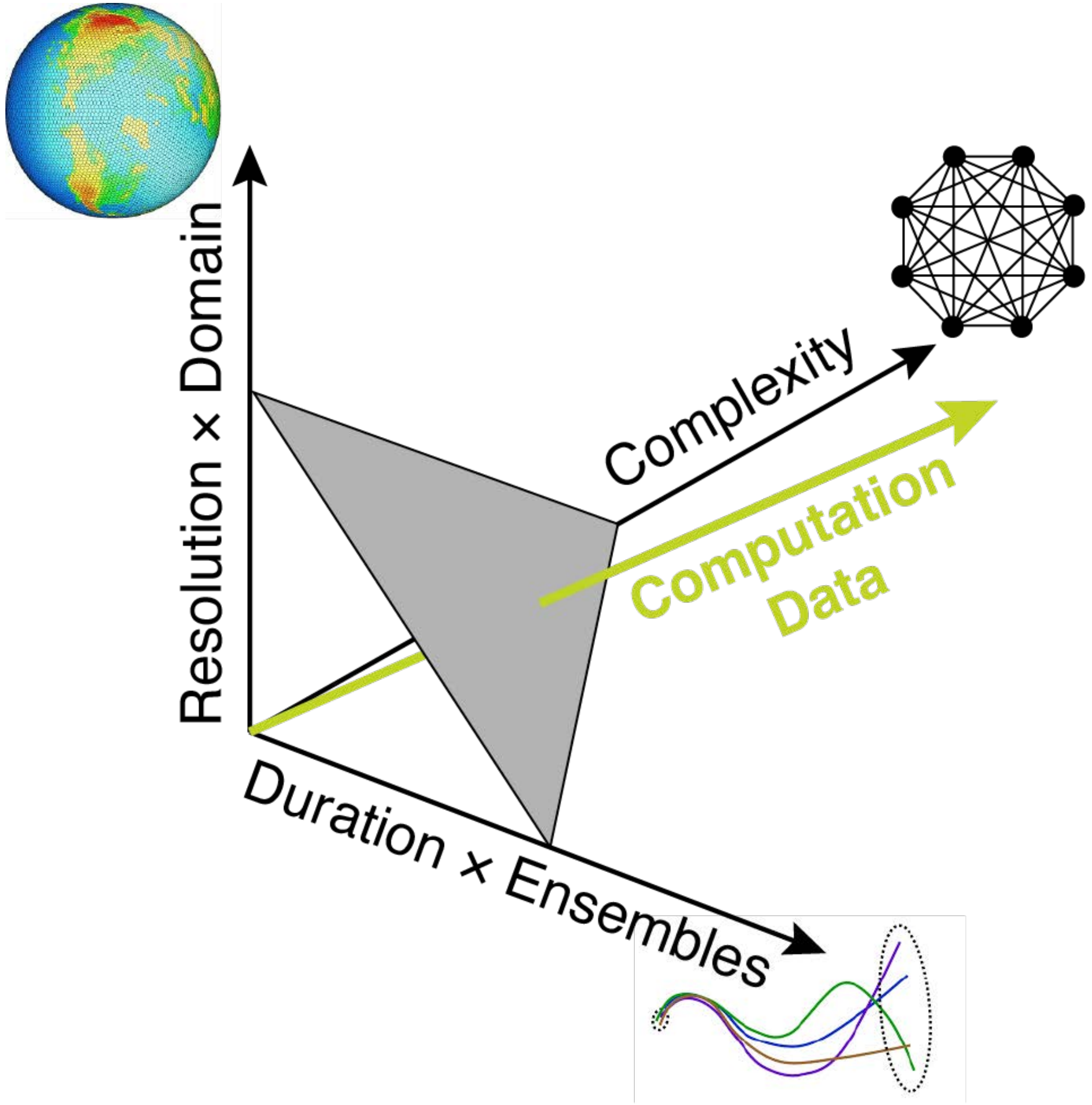
Big Data

Computation for *analysis* is a serious effort

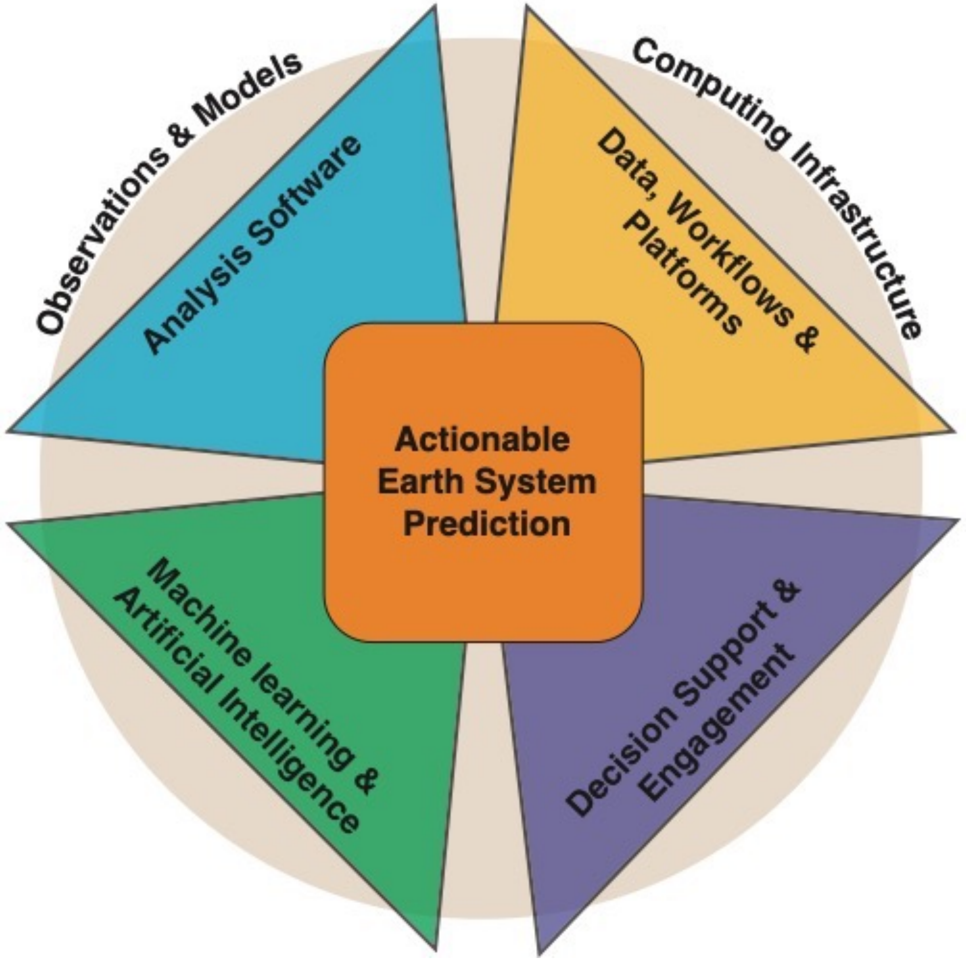
- Collaborative development
- Data science/ML

Decision support “dashboards”

Computational Narratives: Platforms for community & policy engagement



Earth System Data Science



ncar.github.io/esds

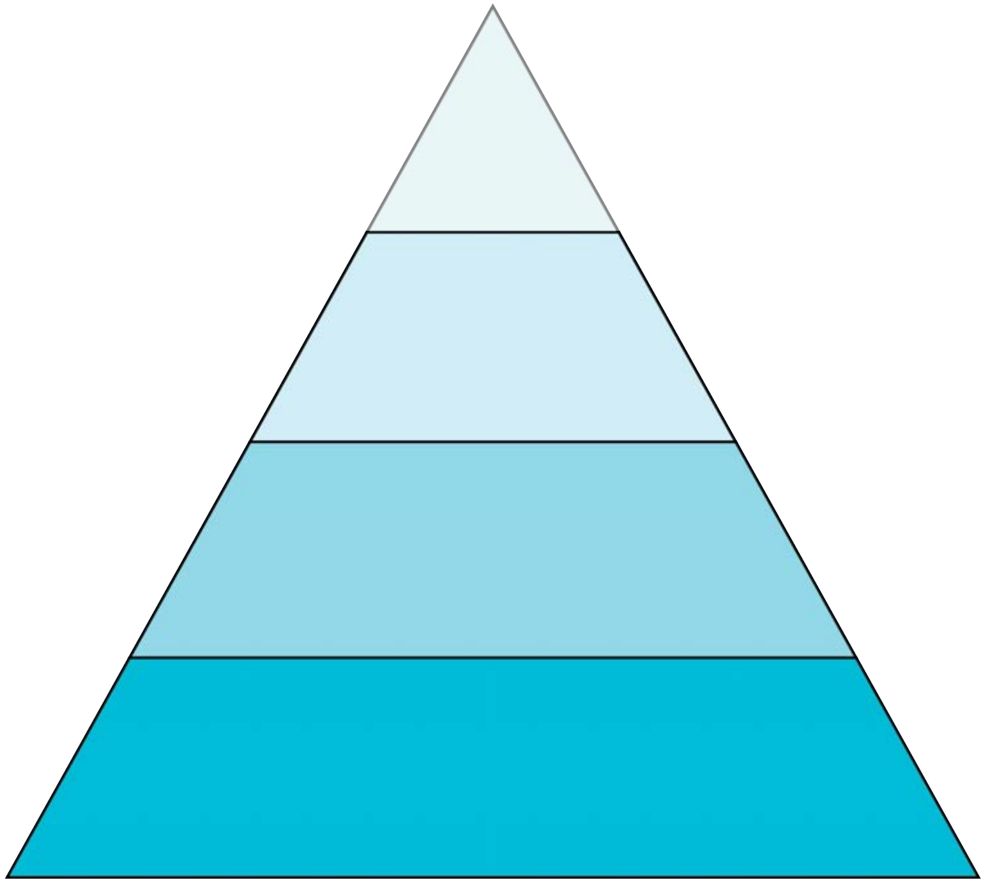
NCAR

PANGEO
pangeo.io

PROJECT PYTHIA
projectpythia.org

LEAP
leap.columbia.edu

Modeling frameworks to support ocean CDR & MRV



CDR processes & experimental framework

Observing system design
Verification with sequential data assimilation

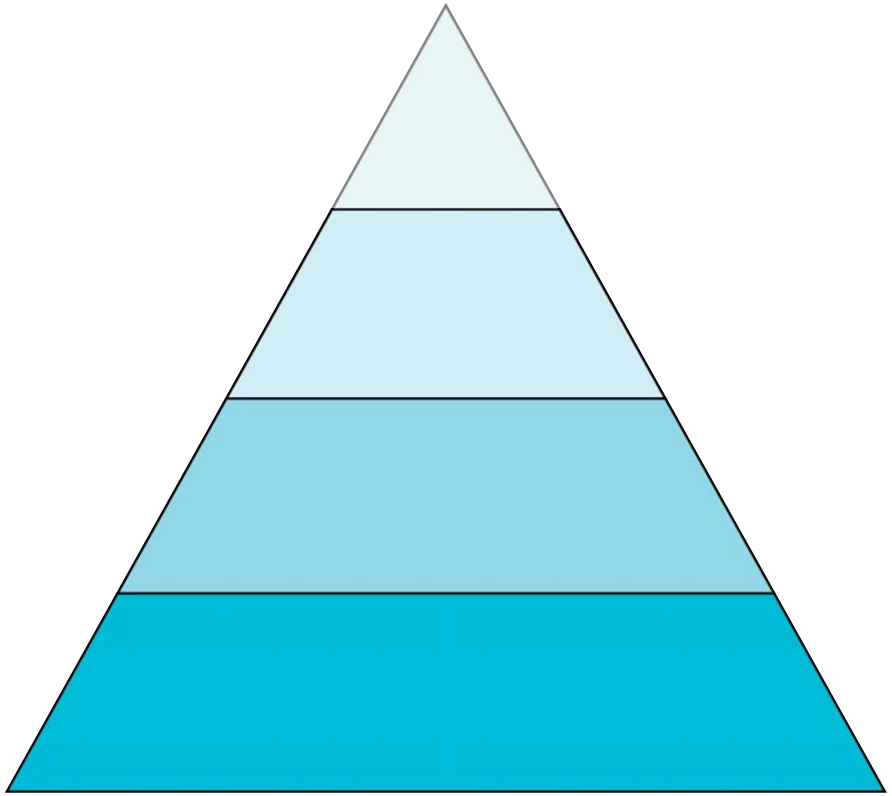
Data assimilation

Ocean physical and biogeochemical models

Proposal

$10M

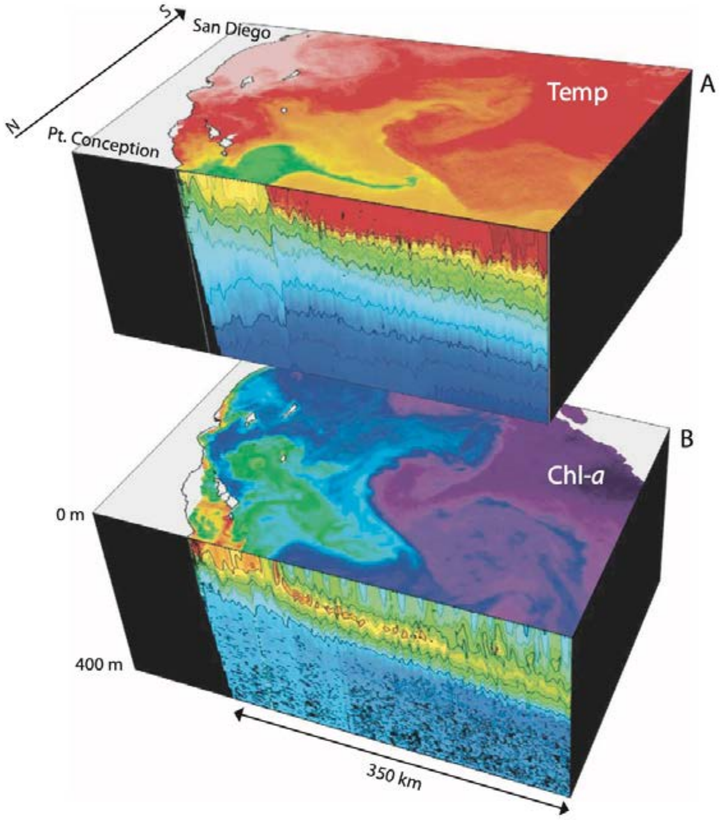
Bespoke modeling applications



+

$10M

Highly resolved observational deployments



Ohman et al. 2013

Thanks!

Contact: mclong@ucar.edu

NCAR

NSF

This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.